REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE 30 Jan 98	3. REPORT TYPE AND DATES COVERED	
4. TITLE AND SUBTITLE A Wavelet Fractal Method for Content Based Image and Video Compression			5. FUNDING NUMBERS
6. AUTHOR(S) Robert J. Bonneau			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Columbia University			8. PERFORMING ORGANIZATION REPORT NUMBER 97-040D
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) THE DEPARTMENT OF THE AIR FORCE AFIT/CIA, BLDG 125 2950 P STREET WPAFB OH 45433			10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES			
12a. DISTRIBUTION AVAILABILITY STATEMENT Unlimited Distribution In Accordance With AFI 35-205/AFIT Sup 1			12b. DISTRIBUTION CODE
13. ABSTRACT *(Maximum 200 words)*			
14. SUBJECT TERMS			15. NUMBER OF PAGES 147
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

Abstract

A Wavelet Fractal Method For Content Based Image and Video Compression

Robert J. Bonneau

Traditional methods of image and video coding rely on linear transformations that focus primarily on high compression. With the increasing demand for digital imagery and video there is now a need for functionality of the compressed information. This dissertation develops a new framework for compression that uses a fractal wavelet method to break the imagery into shape, texture, color, and motion. With this new organization, image information is readily accessible to the user in compressed form. Based on this compression method, we then develop an object-oriented video format.

Image analysts tend to break imagery into the categories of shape, texture, color, and motion. Thus, we begin our approach to image compression by finding mathematical methods that preserve shape and texture in an efficient manner. This new non-traditional method begins by using fractals. A fractal is an object which when observed at its smallest level of detail resembles the overall object itself. Some natural examples include ferns, snowflakes, clouds, and mountains. Recently, engineers have applied fixed point theory to describe a method of fractal image compression . Unfortunately fixed point theory only provides a partial description of fractal compression, since it says little about the spatial frequency structure behind the process.

For insight into the frequency structure of fractal compression, engineers and scientists have recently turned to wavelets, since wavelet methods mirror the fractal compression process. Understanding the frequency structure enables us to see how such a shape and texture work

together in the fractal compression process. With this new insight into fractal compression, we may design a much more efficient compression system. Using this new compression process we also include color and motion. The system developed can then be used for compression of both imagry and video as well as analysis and comparison of the compressed information

Abstract

A Wavelet Fractal Method For Content Based Image and Video Compression

Robert J. Bonneau

Traditional methods of image and video coding rely on linear transformations that focus primarily on high compression. With the increasing demand for digital imagery and video there is now a need for functionality of the compressed information. This dissertation develops a new framework for compression that uses a fractal wavelet method to break the imagery into shape, texture, color, and motion. With this new organization, image information is readily accessible to the user in compressed form. Based on this compression method, we then develop an object-oriented video format.

Image analysts tend to break imagery into the categories of shape, texture, color, and motion. Thus, we begin our approach to image compression by finding mathematical methods that preserve shape and texture in an efficient manner. This new non-traditional method begins by using fractals. A fractal is an object which when observed at its smallest level of detail resembles the overall object itself. Some natural examples include ferns, snowflakes, clouds, and mountains. Recently, engineers have applied fixed point theory to describe a method of fractal image compression . Unfortunately fixed point theory only provides a partial description of fractal compression, since it says little about the spatial frequency structure behind the process.

For insight into the frequency structure of fractal compression, engineers and scientists have recently turned to wavelets, since wavelet methods mirror the fractal compression process. Understanding the frequency structure enables us to see how such a shape and texture work

together in the fractal compression process. With this new insight into fractal compression, we may design a much more efficient compression system. Using this new compression process we also include color and motion. The system developed can then be used for compression of both imagry and video as well as analysis and comparison of the compressed information

A Wavelet Fractal Method for Content Based Image and Video Compression

Robert J. Bonneau

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in the Graduate School of Arts and Sciences

Columbia University
1997

Abstract

A Wavelet Fractal Method For Content Based Image and Video Compression

Robert J. Bonneau

Traditional methods of image and video coding rely on linear transformations that focus primarily on high compression. With the increasing demand for digital imagery and video there is now a need for functionality of the compressed information. This dissertation develops a new framework for compression that uses a fractal wavelet method to break the imagery into shape, texture, color, and motion. With this new organization, image information is readily accessible to the user in compressed form. Based on this compression method, we then develop an object-oriented video format.

Image analysts tend to break imagery into the categories of shape, texture, color, and motion. Thus, we begin our approach to image compression by finding mathematical methods that preserve shape and texture in an efficient manner. This new non-traditional method begins by using fractals. A fractal is an object which when observed at its smallest level of detail resembles the overall object itself. Some natural examples include ferns, snowflakes, clouds, and mountains. Recently, engineers have applied fixed point theory to describe a method of fractal image compression . Unfortunately fixed point theory only provides a partial description of fractal compression, since it says little about the spatial frequency structure behind the process.

For insight into the frequency structure of fractal compression, engineers and scientists have recently turned to wavelets, since wavelet methods mirror the fractal compression process. Understanding the frequency structure enables us to see how such a shape and texture work

together in the fractal compression process. With this new insight into fractal compression, we may design a much more efficient compression system. Using this new compression process we also include color and motion. The system developed can then be used for compression of both imagry and video as well as analysis and comparison of the compressed information

Contents

List of Figures

List of Tables

List of Abbreviations

WT - Wavelet Transform

CWT - Continuous Time Wavelet Transform

DWT - Discrete Wavelet Transform

FT - Fourier Transform

FS - Fourier Series

DFT - Discrete Fourier Transform

STFT - Short Time Fourier Transform

MRA - Multiresolution Analysis

FIR - Finite Impulse Response

IIR - Infinite Impulse Response

DCT - Discrete Cosine Transform

NTSC - National Television Standards Committee

HDTV - High Definition Television

JPEG - Joint Photographic Experts Group

MPEG - Moving Pictures Experts Group

IFS - Iterated Function System

IFSp - Iterated Function System with probabilities

PIFS - partitioned iterated function system

QMF - quadrature mirror filter

ITT - Iterated Transformation Theory

H-space - Hilbert Space

B-space - Banach Space

List of Notations

iff - if and only if

supp(f) - support of $f(x)$

dom(T) - domain of the operator T

range(T) - range of the operator T

N(T) null space of T

r(T) spectral radius of T

$\Re$ - set of all real numbers

$\Re^n$ - set of all real N-tuples $x = (\xi_1, \ldots, \xi_N)$

$\Re_+$ - set of nonnegative real numbers

l^2 - the space of square sumable sequences

$\|\ldots\|$ - norm

$\rho(\ldots, \ldots)$ - metric

$\langle \ldots | \ldots \rangle$ - scalar product

Acknowledgments

I wish to express my gratitude to Professor Henry E. Meadows, my advisor, for his vision and support in guiding me through my thesis. I would also like to thank Dr. Alexandru Bogdan for his patience and enthusiasm in my topic. I would like to thank my committee of Prof. Dimitris Anastassiou, Prof. Daniel Beshers, Prof. Alex Eleftheriadis, Prof. Arlin Crotts for their encouragement through the defense process. I would like to thank all my friends at Columbia - particularly Paul Christianson for his consistent support for this project. I also like to thank Lt. Col. David Dimiduk, Col. Lanny Larson, Mr. Bill Thompson, Dr. Don Washburn, Dr. Jon Sjogren, and all the other Air Force people that made this project possible. I would like to thank my parents, Byron and Marcia Bonneau for their undying love and support over the years in all my academic pursuits. Finally, I would like to thank my fiancee Sonya for her love and support over the last year.

Chapter 1

Introduction

Contents

1.1 Motivation

The speed of information exchange has increased dramatically over the last several decades. This speed has brought with it increased need for better compression of everything from photographs to movies. In order to meet this demand we look to new methods to decompose imagery into an extremely compact form that can be analyzed for its content. Recently international committees have convened to promote the new video and image compression standards. These committees have outlined features in compression that are desirable for new compression methods. We will first review the existing standards and then look at the proposed new directions.

The existing standard has been based on the JPEG image compression format which uses the discrete cosine transform as a fundamental technique of compression. The DCT is a linear transform with excellent compression properties. This compression standard has focused on high compression with accurate reproduction as its main criteria. The MPEG1 and 2 video standards have continued this tradition using block based DCT methods as their foundation with provisions for high resolution, spatial, temporal, and SNR scalablity and a host of other features designed for improvement of speed in encoding and decoding of video.

Recently the MPEG 4 requirements have changed from traditional techniques to less traditional non-DCT based methods. With the growing emphasis on wireless communications, new avenues for compression have been explored such as wavelet, vector quantization, and fractal techniques. These new techniques have promised higher compression ratios with less image distortion. These new methods have also focused on the trend for interactive computer applications where objects in scenes are defined much as a human would describe them. An example might be a car separated from a building separated from a human as individual objects within a scene.

Continuing the trend of object based imagery and video MPEG-7's goals are to establish a new compression format where semantic information such as color, texture, and shape can be easily extracted from image and video in the compressed domain. Such information should be used for identification of objects and their relative position and direction if they are in motion.

1.2 Some Basic Elements of Imagery and Video

Imagery and video as the human visual system interprets them can be broken into 4 basic elements, shape, color, texture, and motion. These elements must be an integral part of the compression process in an object-based interpretive system. Shape has one of the most difficult techniques of pattern recognition because finding an invariant measure of shape is a challenging process. A basic tasks of shape analysis is to find a reliable edge outline of an object. New techniques in wavelet multiresolution analysis are simplifying this process and are making content based shape identification more practical.

Texture is also becoming an increasingly used technique of image segmentation. Texture is in increasingly useful when there is no color information and an absence of well defined shapes in an image. There are many ways that have been used to describe texture from Fourier desciptors to eigenvalue principle component analysis but texture description is very much an open topic. Recently, however, fractals have allowed engineers to develop some fundamental texture descriptors. Combined with new wavelet techniques fractals are an increasingly popular means of texture discrimination.

Another often used method of image analysis is image color image segmentation. Most color imagery can be segmented into 3 primary colors R, G, and B or some related color set. Thus by seeing the relative colors present in a given region of the image a user can identify objects that

have the same percentage of these three colors.

Finally, motion is a final issue when for moving image segmentation and analysis. Many methods have been developed recently to interpret motion within successive frames of video. In order to put this motion analysis into the context of object based video compression, we must associate motion vectors with individual objects. We therefore explore a new way of performing such an operation with a wavelet-based optical flow technique .

1.3 Encoder/Decoder Design

Speed for both image encoders and decoders is essential. Even with improvements in computer hardware speed the demands placed on image coders and decoders are increasing proportionally. The demands for real time compression and decompression are placing stress on coder design; thus whatever algorithm is used to code data must be computationally simple and easy to implement on existing real time hardware. For decompression, any user of a given decoder should be able to decompress a video sequence on a normal personal computer with no special decompression hardware. Obviously, the compression process should also not significantly distort the original image or video sequence.

Scalability [52] refers to an image format's ability to adapt to different bandwidth and speed capabilities of different video systems. For instance, some video users may be connected directly to the internet through high bandwidth fiber connections and do not need to worry about bandwidth, but are more concerned with image quality. Some users however simply want to connect to a video player or conferencing over their phone line and are willing to sacrifice image quality for low bit rate. Spatial scalability refers to the ability to select the spatial resolution of an image in

accordance with bandwidth and speed requirements. Progressive JPEG, many wavelet techniques, and fractal encoding share this ability to present the user with a low resolution copy of an image and then progressively add detail to this representation when the higher resolution information becomes available over the communication channel.

Temporal scalability is another method of controlling data flow. Temporal scalability allows the user of a video compression system to adaptively select the number of frames transmitted within a given time interval to match speed and bandwidth requirements. For instance, normal NTSC video runs at 30 frames per second or 60 interlace fields per second. Below 30 frames per second the human eye can detect visible gaps in motion within a video sequence but for some high speed film applications higher than 30 frames per second is necessary. Thus depending on the customers needs, the rate of the video can be scaled accordingly.

With the advent of MPEG-4 a new scalability has emerged based on video objects . Video objects are used in conjunction with the concept of motion compensated video where objects in a scene are transmitted in the first frame and only their motion is transmitted in sucessive frames until the object significantly changes shape or disappears from the scene. As pattern recognition principles improve this technique will continue to advance.

1.4 Overview

To accommodate these analysis and compression needs we turn to the concept of the fractal. The fractal has emerged to describe many different natural phenomenon. In the work on turbulence theory Mandelbrot [46] developed the concept of a fractal to describe an object which is self similar at any spatial scale at a which you observe it.The essential characteristic of fractals is that the smallest structure of the fractal resembles the overall

Figure 1-1 Fractal Fern

structure. This evidenced in the fractal fern of Figure 1-1. This concept of the fractal and fractal image compression has been mathematically described through linear approximation theory. Barnsely[11], Jaquin[37], Bogdan[2] and others have used the concept of metric space, the iterated function system, and attractors to describe the process of fractal or self vector quantization compression. Unfortunately linear approximation theory does not sufficiently reveal the frequency structure of fractals or fractal compression.

Recently, others such as Davis[20], Rinaldo and Calvagano[57] have shown that there is a direct analog of fractal compression frequency segmentation in wavelet compression. This realization is extremely valuable since wavelet signal processing and compression techniques are a well understood signal processing method. In parallel other researchers such as Arneodo and Bacry[1], as well as Mallat[45] have demonstrated in a 1-dimensional sense how wavelets can be used to reveal the underlying structure of the fractal and guide a linear approximation approach for synthesis of a wavelet decomposed signal. With this new insight into the nature of fractals and fractal compression we can develop a new method of fractal image compression combining both fractal and wavelet methods.

This new fractal-wavelet method has the ability to break an image apart into the basic elements of shape and texture. A basic way to define shape of an object is in terms of its geometrical outline. Recently Mallat[44] has developed a wavelet technique which parallels an earlier method developed by Canny [14] to reveal consistent outlines or shape of objects in two dimensions. On the other hand, fractal dimension has been used to describe the two dimensional textural regions in imagery[54]. Recently Flandrin[28] the wavelet model to reveal an objects fractal dimension by the decay of energy across frequency bands.

What we find through the combination of the linear approximation and wavelet definitions of shape and texture, is that there is a direct connection between two. Logically this connection makes sense in light of the concept of a fractal since, in a fractal, the smallest part resembles the overall shape of an object. The fractal process thus defines a mapping between the overall object and its smallest components. Thus the shape of the object at a smaller scale *is* the object's texture. With this fact in mind we can develop a new concept for image compression that combines shape and texture models not only for analysis but as an essential part of the compression process. Thus

compressed information can be analyzed without decompressing because it is in a form geared toward analysis. We will then extend this principle to color imagery and video.

Chapter 2

Mathematical Basis of Fractal Image Coding

Contents

2.1 Metric Spaces

To develop the basic technique of fractal compression we must find a mechanism by which an image can be decomposed and reassembled or mathematical terms analyzed and synthesized. Do show how this mechanism works we must first provide an overview of the concept of a space. The concept of space defines a set of rules for mapping a region of an image onto itself in a reproducible way. This mapping defines the essential structure of the fractal or the fractally compressed object.

A space is a set between whose elements certain relations are prescribed by means of axioms; the set is said to have been given the structure of the relevant space. [4] At the very specialized space level we have Hilbert space (H-space) which is a special case of Banach space (B-space), which is a linear space and complete metric space, which is an important subset of more general class of topological spaces. Figure 2-1 shows the relationships between spaces. A H-space is an immediate generalization of R^N while a topological space is at a higher level the hierarchy and can be seen as a generalization of the concept of an open set in R^N. A (linear) vector space V is a nonempty set of elements called vectors for which we define two algebraic operations: addition of vectors and multiplication of a vector by a scalar (number). A finite set of elements $\{f_i \in V: i = 1,.....,n\}$ is independent if a set of scalars $\{\lambda_i\}$ such that $\lambda_1 f_1 + ...\lambda_n f_n = 0$ cannot be found. A finite set $\mathbf{S} \subset \mathbf{V}$ is a basis in V if it is linearly independent and any vector V is a linear combination of elements in S.

Definition 2.0 *A set* **X** *is called a* **metric space** *if for each pair of elements f,* $g \in \mathbf{X}$ *there is associated a non-negative real number* $\rho(f,g)$*, the distance between f and g subject to the following conditions:*

1. $$\rho(f, g) \geq 0; \ \rho(f, g) = 0 \ \ iff \ \ f = g$$

2. $$\rho(f, g) = \rho(g, f)$$

3. $\rho(f, g) \leq \rho(f, h) + \rho(h, g)$ for any h $\in \mathbf{X}$ (the triangle inequality)γ

2.1.1 Banach Space

We add analytic properties to a vector [4] space V when we define a norm (generalized absolute value). A norm $\|\cdot\|: V \to \Re_+$ associates with every element in **V** a nonnegative real number. A *normed space* (a vector space having a fixed norm) is metrizable if we set:

$$\rho(f, g) = \|f - g\| \tag{2.1.1}$$

for all $(f, g \in \mathbf{V})\gamma$. We note that only the distance functions for which $\rho(f,g) = \rho(f-g,0)$ will introduce a norm. Very important in applications are the complete normed spaces which are called Banach spaces (B-Space). In a B-space the notions of completeness and convergence take the

form of convergence in norm. A sequence $\{f_n\}$ in B-space converges iff. there exists an element

$f \in \mathbf{V}$ such that

$$\lim_{n \to \infty} \|f_n - f\| = 0 \tag{2.1.2}$$

2.1.2 Hilbert Space

We can add geometric properties such as orthogonality [4] between two elements in linear space

through a *scalar product* (inner product).

Definition 2.1 A **scalar product** (inner product) (.l.) on a linear space V over a filed of scalars

κ is a function (f,g): $\mathbf{V} \times \mathbf{V} \to \kappa$ such that the following conditions are satisfied for all

$(f, \dot{g}, h \in \mathbf{V})$ and all $\alpha, \beta \in \kappa$

1. $\quad . (f, \alpha g + \beta h) = \alpha(f|g) = \beta(f|h$

2. $\quad (f|g) = (g|f)$

3. $\quad (f|f) > 0 \; iff \; f \neq C$

A pre-Hilbert space is a vector space **V** with scalar product. This space becomes a normed space if

we define a norm on **V** through.

$$\|f\| \stackrel{\text{def}}{=} (f|f)^{\frac{1}{2}}$$

If this space is a B-space then we call **V** a *Hilbert space* (H-space).

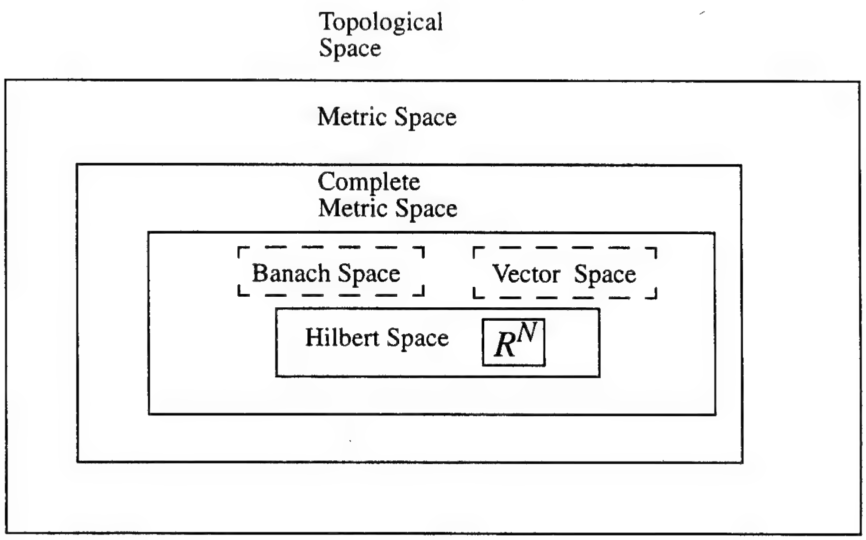


Figure 2-1 Diagram of Spaces

2.2 Fixed points

Fixed points are now the beginnings of a method by which we can transform a space into a desired function or the basic principle in the mechanism of constructing a fractal. This approach begins with the method of successive approximations. We want to apply the sucessive approximation method in the form $x_{n+1} = Tx_{n+1}$ First we introduce contractive operators on a metric space and then we present the main result

Definition 2.3 *An operator* $T:\Omega \subseteq \Omega \to \Omega$ *on a metric space* (X,ρ) *is called s contractive iff for*

$0 \le s < 1$ $\rho(Tx, Ty) \le s\rho(x, y)$ *for all* $x, y \in \Omega$. *If* $s = 1$, *T is called* **nonexpansive**; *if the above equation holds for* $0 < s < \infty$, *T is called* **Lipschitz continuous.** *If* $\rho(Tx, Ty) \le \rho(x, y)$ *for all* $x, y \in \Omega$ *with* $x \neq y$ *then T is called* **contractive.**

Theorem 2.0 (Banach 1922) Given an s-contractive operator $T:\Omega \subseteq \Omega \rightarrow \Omega$ where Ω is closed and non-empty in the metric space (X,ρ), the following are true:

1. Existence and uniqueness: T has exactly one fixed-point on Ω or equivalently or equivalently equation 4 has a unique solution.

2. Convergence of iteration: the sequence $\{x_n\}$ of sucessive approximations convergence to the solution x for any choice of the initial approximation x_0.

3. Error estimates: for all n = 0,1,2,... we have the a priori error estimate

$$\rho(x_n, x) \le \frac{s^n}{(1-s)}\rho(x_0, x_1) \tag{2.2.1}$$

and the a posteriori error estimate

$$\rho(x_{n+1}, x) \le \frac{s^n}{(1-s)}\rho(x_0, x_1) \tag{2.2.2}$$

4. Rate of convergence: for all n = 0,1,2,... we have

$$\rho(x_{n+1}, x) \le s\rho(x_n, x) \tag{2.2.3}$$

To apply the fixed point principle and the contraction mapping theorem (2.0), we need a slightly different setting. Let X be a complete metric space and H(X) be the set of nonempty compact subsets of **X**. The distance between a point x and a set $A \subset \mathbf{X}$ is

$$\rho(x, \mathbf{A}) = \inf_{a}\{\rho(x, a):\ a \in \mathbf{A}\} \tag{2.2.4}$$

The Hausdorf distance is defined for $\mathbf{A},\mathbf{B} \in H(\mathbf{X})$ by

$\rho(\mathbf{A},\mathbf{B}) = \sup\{\rho(a,\mathbf{B}), \rho(b,\mathbf{A}) : a \in \mathbf{A}, b \in \mathbf{B}\}$ It can be shown that (H(**X**), ρ_h) is a complete metric space and we can use Banach's fixed point fixed point theorem (2.0). Let's define W = $\{w_n, \ldots\ldots, w_N\}$ a finite collection of contraction mappings in **X** and let $W(K) = \cup\, w_i(K)$, $\mathbf{K} \in \mathbf{X}$

It follows that W is a contraction map on H(**X**) and we can use the contraction mapping principle. This result can be reformulated on the metric space of **X**.

Theorem 2.0(Hutchinson). Let $(\mathbf{X},\rho)$ be a complete metric space and $W = \{w_1, \ldots\ldots, w_N\}$ be a finite set of contraction maps on **X**. Then there exists a unique closed bounded set **K** such that $K = \cup_{i=1}^{N} w_i(K)$. Furthermore, **K** is compact and is the closure of the set of fixed points $k_{i_1, \ldots\ldots\ldots, i_p}$ of the finite compositions $w_{i_1} \circ \ldots\ldots\ldots \circ w_{i_p}$. Starting with an arbitrary set $\mathbf{A} \subset \mathbf{X}$ the iterative method $W^n(\mathbf{A}) \rightarrow \mathbf{K}$ converges in the Hausdorff metric.

2.2.1 Iterated Function Systems

Iterated function systems is the discrete or finite mathematical mechanism for the[4,11] the

synthesis of a fractal. There two types of IFS, global and local. We will first discuss the more general global IFS and then move on to local IFS.

The collection $W = \{w_1,.....,w_N\}$ is know as an iterated function system (IFS). In binary and some new forms of greyscale image coding we take $\mathbf{X} = \mathbf{A} \subset \Re^2$ a rectangle and let $\mathbf{K} = \mathbf{I}$ be the image pattern. Usually the attractor **K** has fractal properties like detail at every scale. It turns out that images which look very complicated can be represented with a very small number of transformations.

One of the most important questions is if we can build a fractal code that approximates well enough a given image or signal. One approach is by taking a fractal code T by construction, starting with the assumption the original image *f* is a fixed-point $f = T(f) = \sum T_{ij}(f)$ Because we limit the possible maps T and we do not know the fixed point T beforehand, we are looking for some map such that $\rho(f, Tf) \leq \varepsilon$. A bound to the distance of the fixed-point to the original can be estimated if T is the contraction. As a consequence of the contraction mapping principle 2.0 we have the collage theorem that can be reformulated in the metric space (I, ρ).

Theorem 2.1 In the metric space (I, ρ) of images with the same support let the fractal code $T \in T \subset B(I, I)$ be a contraction with $s_T < 1$. Then we have the following upper bound of the distance between the fixed-point $f^* = T(f^*)$ and the original image $f \in I$.

$$(f, f^*) \leq \frac{1}{1 - s_T}\rho(f, T(f)) \tag{2.2.5}$$

2.2.1.1 The Global IFS

Recall that an [4,11]iterated function system (IFS) is a collection of transformations $\{w_1,......,w_n\}$ from a metric space $(\mathbf{X},\rho)$ into itself. If $W(K) = \bigcup_i w_i(\mathbf{K}), \mathbf{K} \in \mathbf{X}$ is a contraction on the metric space $(\mathbf{X},\rho)$, then iterating W starting with any element in the space converges to the unique limit **I.** Such an IFS is known as a global IFS since it operates on the entire metric space. An IFS can encode black and white images when I is a subset of R^2.

A random iterated function system, also called an (IFSp) is an IFS together with a probability vector $\{p_1,......,p_n\}$. A contractive IFSp determined by $\{X,w_i,p_n\}^n{}_1$ will converge to a unique set $\mathbf{I} \in \mathbf{X}$ which is the support of the associated invariant measure. Invariant measures generated by dynamical systems are studied in ergodic theory and will be discussed in Chapter 4 in detail. A gray scale image can be modeled as the natural measure of an IFSp and can be generated using the chaos game. Consider the attractor $\mathbf{I} \subset \mathbf{A}$, where A is a bounded subset of R^2 usually a rectangle. Then I is an element of *H(***A***)* together with the Hausdorf distance $\rho(.,.)$ is a complete metric space under certain conditions. To generate the attractor **I** we start with any point $z_0 \in \mathbf{I}$ then all the points z_i will be the attractor, else after a number of iterations, z_i will be the attractor. In practical implementations, if we do not know the attractor, we can discard the first iteration. The gray image appears as the natural measure μ generated by a dynamical system we have described. Let *M(***A***)* be the space of all Borel probability measures on **A** and let **B** be a Borel subset on **A.** Then

$$\mu(\mathbf{B}) = \frac{\sum_{i=1}^{\infty} \mathbf{1}_B(z_i)}{\sum_{i=1}^{\infty} \mathbf{1}_A(z_i)} \tag{2.2.6}$$

where $\mathbf{1}_B(z_i)$ is the indicator function for the set **B**. To visualize the measure on a computer display, we have to compute $\mu(\mathbf{B}_i)$ for each pixel and rescale the dynamic values to match the dynamic range of that display. Now we can rewrite the fixed-point equation directly in the space of probability measures.

Under some technical conditions, $M(\mathbf{A})$ endowed with the Hutchinson metric is a complete metric space and we can define the Markov operator which is the contraction:

$$T(\mu)(\mathbf{B}) = \sum_{i=1}^{n} p_i \mu\left(w_i^{-1}(\mathbf{B})\right) \tag{2.2.7}$$

The contractivity of $T(\mu)$ was proved for the hyperbolic case when all $\{w_i\}$ are contractions. We have seen that $\{w_i\}_1^n$ uniquely determines the attractor **I** of the support of the IFSp. For different sets of probabilities $\{p_i\}$ we obtain different gray level images, all having the same support. Several IFS's can be combined to obtain a complicated image. Next we look at a special type of IFSp where the transformations $\{w_i\}$ are restricted to a particular region of the space **A.**

2.2.1.2 The Local IFS

The local iterated function system was named fractal block coding by Jacquin[37], local interated function system (LIFS) and the fractal transform by Barnsely[12] or partitioned iterated

function system (PIFS) by Fischer, Jacobs, and Boss[27]. It is the basis of what is currently known as fractal image compression. The support of the gray scale image is a bounded rectangle $\mathbf{I} \in R^2$ which is invariant under a block-wise continuous transformation $W:R^2 \to R^2$. In a *PIFS* we obtain more flexibility by having the domain of each w_i restricted to a subset $\mathbf{D}_i \subset \mathbf{I}$. Note we can find an infinite number of combinations $\{w_i\}$ with the same invariant set **I.** Also *W* may have several invariant sets.

A grey level image is modeled as the invariant probability measure μ generated by the iteration of *W* on the complete metric space *H*(A) and the associated mapping *T* given by $\{w_i, p_i\}$ operating on the space of probability measures *M* with the support **A** as before $T:M(\mathbf{A}) \to M(\mathbf{A})$ Since in PIFS coding, the invariant set of the *IFS* is the space **A,** we have $\mathrm{I} = \mathrm{A} = supp(\mu)$,

Contractivity of each T_i is given by $p_i < 1$ and guarantees convergence of the iteration of T to limit μ^* in the space of probability measures. The limit μ^* is the approximation of the image μ to be coded, and T is called the fractal code for μ. In a practical coding system, I is partitioned in rectangular blocks b_{ij}. Block mapping w_{ij} and a massic (greytone) transformation with a gain factor p_{ij} for each block such that

$$T(\mu) = \sum T_{ij}(\mu) = \sum p_{ij} * \mu(w_{ij}(b_{ij})) \approx \mu \tag{2.2.8}$$

Coding the image μ is equivalent to finding all mappings T_{ij}, the restriction of T to each block of the partition. For decoding we use the chaos game as before, or we can apply the Markov operator T in a successive approximation algorithm (start with any initial image $v \in M$)

$$\mu^* = (T)^n(v), n \to \infty \tag{2.2.9}$$

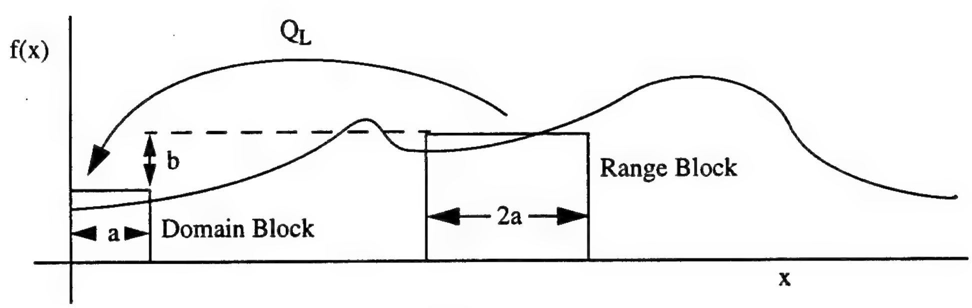
The problem of finding $T = \sum T_{ij}$ makes ITT coding very complex and thus we seek an new means of determining the transformation T.

2.3 Fractal Compression Techniques

Extending the concept of the IFS to practical application there are two primary methods that have emerged. Jaquin's Algorithm[37] was based on self vector quantization techniques where the codebook is taken from the image itself. Yuval Fisher [27]extended this technique for a variable partitioning of the image based of feature size and resolution.

2.3.1 Jaquin's Algorithm

Thus we develop what is commonly known as the fractal block transform [37] shown below in Figure 4a



(4a)

2-2 Jaquin's 1-D Block Mapping

This process thus exploits the self-similarity of two spatial scales to compress an image. The process of fractal encoding is thereby described by equation 2.3.1[3]

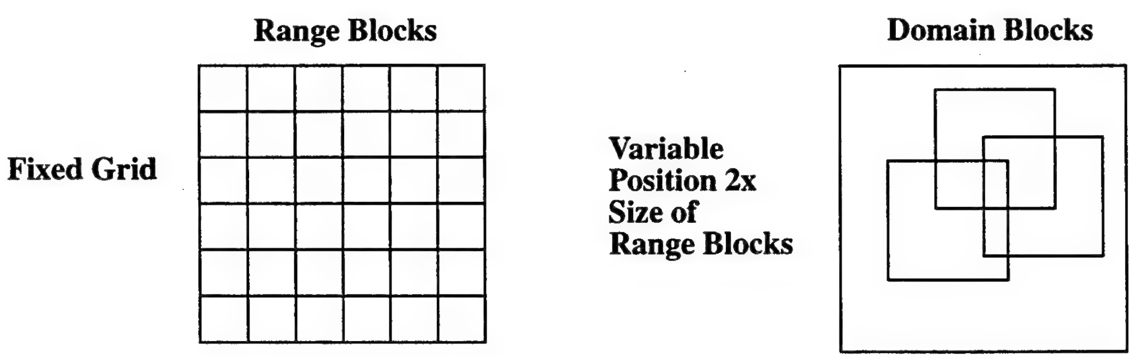
:

$$f(x) = Tf(x) = U_L f(x) + b = Q_L f(2x) + b \tag{2.3.1}$$

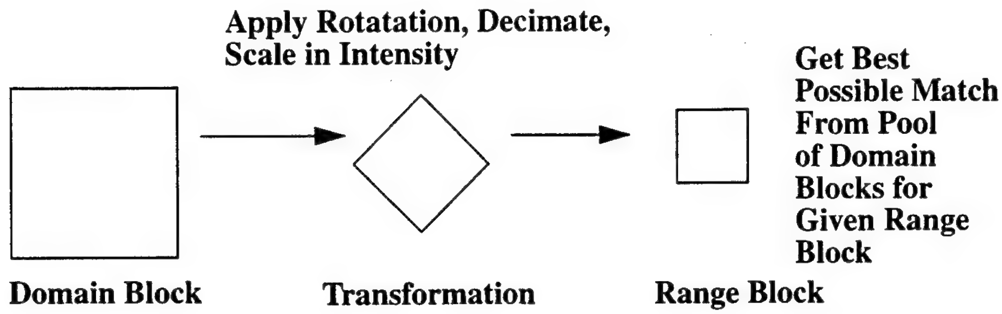
where f(x) is the image to be transformed and T is a contractive operator with unique fixed point $\bar{f}$. Encoding f means finding and operator T having a fixed point $\bar{f}$ approximately equal to f while decoding is equivalent to finding the fixed point $\bar{f}$ by iterating T starting with an image selected at random. U_L represents the transformation applied to the domain blocks which both grey levels and decimates; Q_L represents a simpler U_L that simply scales the already decimated domain blocks, and b is the intensity offset applied to the domain blocks. Note that f, b ε V where V is a discrete and finite dimensional space.

In two dimensions the fractal encoding processes can be described in 4 steps. Generically, fractal encoding involves taking a block of pixels in a image known as a domain block, spatially subsampling the block, and matching it to a smaller block of pixels known as a range block.

Step 1: Divide Image Into Range and Domain Blocks



Step2: Match Domain Blocks to Range Blocks



Step 3: Build a Table of Domain Block Positions, Rotations & Scalings - This is the Compressed Image

Step 4: Reverse Process to Rebuild Image Start with blank image with all pixels set to some arbitrary value and iterate on transformations from coded table.

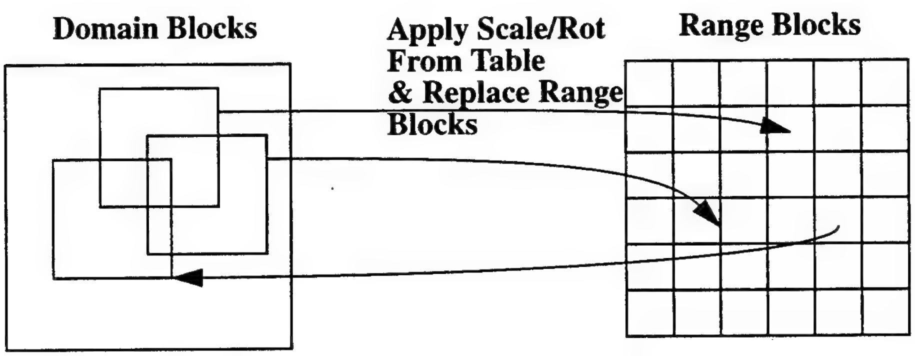


Figure 2-2 Jaquin's 2-D Block Mapping

As is shown in Figure 2-3 Bogdan[2] and Davis[20] showed the fractal block mapping technique takes low frequency coefficients and maps them to high frequency coefficients. Thus the block mapping uses the low frequency information to approximate the high frequency information.

Jaquins Fractal Frequency Mapping

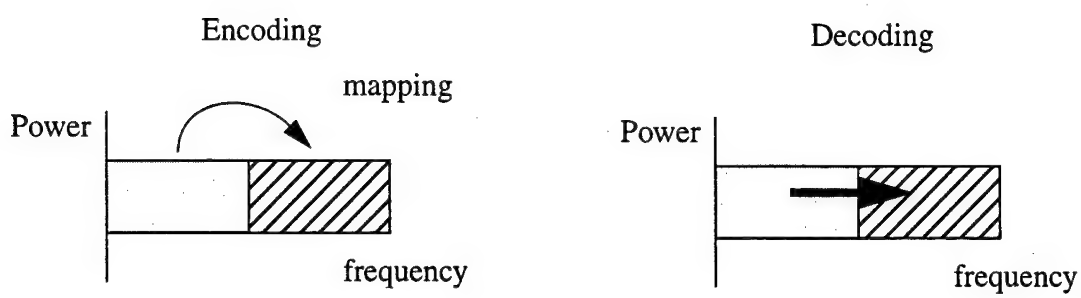
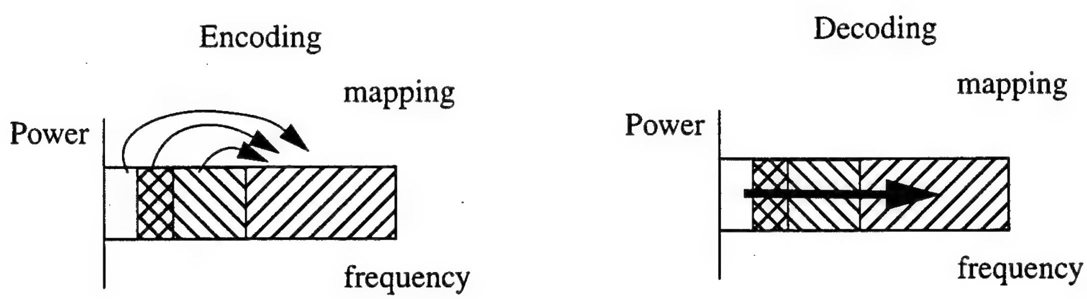


Figure 2-3 Frequency Analysis of Jaquin's Technique

2.3.2 Fischer's Algorithm

The problem with Jaquin's technique is that once a range and domain block size is chosen, all features, large or small must be represented with those block sizes. To remedy this problem, Yuval Fischer [26] has modified the original fractal transform by incorporating adaptive block sizing depending on the amount of information within each fractal range and domain block. This process gives better overall results simply because it preserves spatial detail in those blocks where information is needed and achieves high compression for those regions which are relatively homogenous. Fischer uses different tiling techniques such as quadtree partitioning where blocks remain square as in Jaquin's algorithm and rectangular quadtree which gives greater adaptability to spatial features. Also, Fischer as made use of triangular range and domain blocks.

Fischer's Fractal Frequency Mapping



2-4 Frequency Analysis of Fischer's Technique

As is shown in the Figure 2-4, like Jaquin's technique, Fischer uses the low frequency information to reconstruct the high frequency information. Unlike Jaquin however, Fischer reconstructs with multiple frequency scales from low to high frequency[20]. The width of these frequency scales for the quadtree method is powers of two. Looking at the fractal compression process, we will shown in Chapter 3 a similarity to the wavelet transform frequency breakdown.

Chapter 3

Wavelet Image Coding and Parallels to Fractal Image Coding

Contents

3. 1 Wavelets

As stated at the end of Chapter 2, researchers such as Davis[20] have seen a direct similarity between fractal and wavelet compression. Thus, we are motivated to study the wavelet compression methods from both a 1 and 2 dimensional perspective to gain a better understanding of fractal image compression and develop a means of improving compression performance. First we will describe the general properties of wavelets. We will then introduce two types of wavelets, the QMF Haar and the Gaussian Derivative Spline. We will show how the frequency structure of

the Haar wavelet mirrors that of Jaquin and Fischer's fractal compression and then show how the gaussian derivative Spline has desirable properties which will be used to improve on the fractal compression process. Finally we will give two examples of conventional wavelet image coding techniques for later comparison with the new fractal wavelet method.

3.1.1 Definition

Practically speaking, a wavelet is a means of approximating a function with a set of linear vectors. This set of vectors forms a basis set. This basis is orthogonal if the scalar product of the vectors is equal to zero. Vectors are formed from dilating and translating of some mother wavelet vector $\psi(x)$ which is based on some scaling function $\phi(x)$. The wavelet decomposition thus forms a linear transform. This linear transform can be formulated in terms of linear filters $\psi(x)$ typically corresponds to a highpass filter g(k) $\phi(x)$ typically corresponds to a lowpass filter h(k).

The only constraint imposed on a wavelet function $\psi(x)$ real or complex valued, [25] in order to be a wavelet is the admissibility condition that requires: i f $\psi(x)$ is integrable this actually implies that

$$\int_{R^n} \psi(\mathbf{x}) d^n\mathbf{x} = 0 \text{ or } \psi(|\mathbf{k}| = 0) = 0 \tag{3.1.1}$$

3.1.1.1 Properties

The properties of the wavelet transform, admissibility, similarity, invertibility, regularity, and moment cancellations are described as follows:

Admissibility: For an integrable function, this means that its average is zero.

Similarity: The scale decomposition should be obtained by the translation and dilation of only one "mother" function.

Invertibility: There should be at least one reconstruction formula for recovering the signal exactly from its wavelet coefficients and for allowing the computation of energy or other invariants directly from them.

Regularity: In practice the wavelet should also be well localized on both sides of the Fourier transform, namely it should be concentrated on some finite spatial domain and be sufficiently regular.

Cancellations For some applications, in particular, turbulent signal analysis, the wavelet should not only be of zero value(admissiblity). but should also have some vanishing high-order moments.

3.1.1.2 Analysis

The analysis of a signal involves decomposing with a set of functions which are derived from some common "mother" function. From the function ψ the so called mother wavelet, [25]we generate the family of continuously translated, dilated and rotated wavelets:

$$\psi_{lx'\theta}(\mathbf{x}) = \mathbf{l}^{-\frac{n}{2}}\psi\left[\Omega^{-1}(\theta)\left(\frac{\mathbf{x}-\mathbf{x}'}{\mathbf{l}}\right)\right] \tag{3.1.2}$$

with $l \in \mathbf{R}^+$ as the scale dilation parameter corresponding to the width of the wavelet and $x' \in \mathbf{R}^n$ as the translation parameter corresponding to the position of the wavelet; l and x' are dimensionless variables. l denotes the wavelet scale because it corresponds to the length scale at which analyze f(x) and the position of the analyzing wavelet x', because it indeed corresponds to the actual position in physical space; we must also distinguish x' and x which will be used an integration variable. The rotation matrix Ω belongs to of rotations in R^n and depends on the n(n-1) /

2 Euler angles θ. The factor $l^{-\frac{n}{2}}$ is a normalization which causes all the wavelets to have the same L^2 norm. therefore all wavelets will have the same energy and the wavelet coefficients will correspond to energy densities.

3.1.1.3 Synthesis

Reconstruction of the original signal is known as synthesis. The admissibility condition implies the existence of a reproducing kernel. We can therefore recover the signal f(x) from its wavelet coefficients as shown in equation 3.1.5.

$$f(\mathbf{x}) = C_{\psi}^{-1}\int_{-\infty}^{\infty}\int_{R^n}\int \tilde{f}(l, x, \theta)\psi_{lx'\theta}(\mathbf{x})\frac{dld^n x'}{l^{n+1}}d\mu(\theta) \qquad (3.1.3)$$

3.1.1.4 Comparison With The Fourier Transform

The primary advantage of the wavelet transform is its ability to [72] localize in time or frequency where the fourier transform has great localization in frequency but poor localization in time. Given an absolutely integrable function *f(t),* its Fourier Transform is defined by

$$F(w) = \int_{-\infty}^{\infty} f(t)e^{-j\omega t}dt \qquad (3.1.4)$$

Likewise the discrete-time Fourier Transform is Defined as

$$F[k] = \sum_{n=0}^{N-1} f[n]e^{-j2\pi nk/N} \qquad (3.1.5)$$

The time and frequency localization properties of the wavelet transform can be shown in the following Figure 3-1 using a space frequency analysis plot

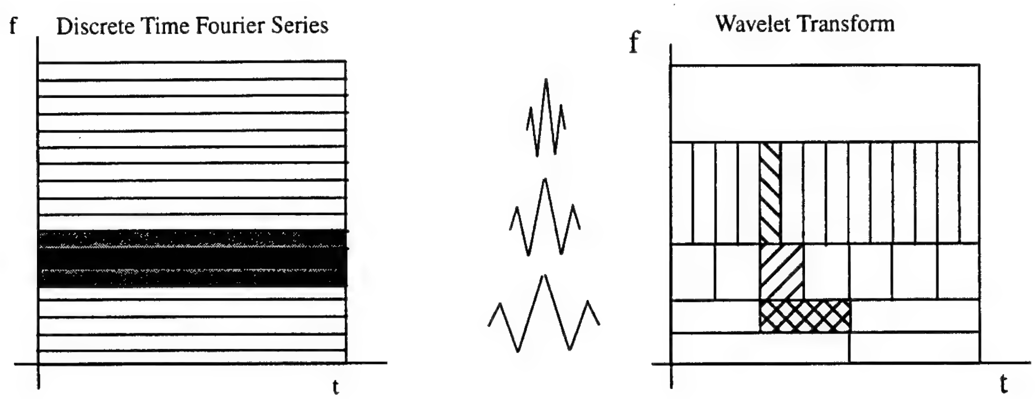


Figure 3-1 Space vs. Frequency Plots of Fourier vs. Wavelet Transform

And finally the short time Fourier Transform is defined by

$$STFT_f(\omega, t) = \int_{-\infty}^{\infty} \omega^*(t-\tau) f(t) e^{-j\omega t} dt \tag{3.1.6}$$

Which like the wavelet transform is localized in frequency because it is windowed in time.

3.1.2 Bases

The expansions of a mother wavelet are described as the [72] bases of the wavelet. Such expansions can be described as orthogonal or biorthogonal. The elements of $f, g \in \mathbf{H}$ where **H**is a H-space, are said to be orthogonal if $(f|g) = 0$. A system $\{f_i, g_j\}$ constitutes a par of biorthogonal bases of a Hilbert space H if and only if

(a) For all i,j in Z (integers)

$$(f_i|f_j) = \delta[i-j] \tag{3.1.7}$$

(b) There exists strictly positive constants A_0, A_1, B_0, B_1 such that for all y in H

$$A_0\|y\|^2 \le \sum_k |(f_k,y)| < B_0\|y\|^2 \tag{3.1.8a}$$

$$A_1\|y\|^2 \le \sum_k |(g_k,y)| < B_1\|y\|^2 \tag{3.1.8b}$$

Let Ω be a closed subset in a complete metric space **X**. A fixed point of a given mapping $T:\Omega \to \Omega$, not necessarily linear is every solution $x \in \Omega$ of the equation 4 $x = Tx$

3.1.3 Moments

At this point it is important to note that the number of [72] vanishing moments of the wavelet is n where

$$0 = \int_{-\infty}^{\infty} t^n \psi(t)dt \tag{3.1.9}$$

Information about the number of moments in a wavelet function can be critical to determining certain frequency properties of the wavelet transform such as Alpha and Holder exponents. This fact will be discussed in detail later in Chapter 4.

3.2 Orthogonal Structures: QMF Haar

To begin with a specific wavelet basis we describe the [73]Haar Wavelet. The Haar Wavelet is one of the most basic wavelet basis and as we will show later can be used to explain the basics of fractal compression since the averaging of a domain block to obtain a range block is simply applying the Haar lowpass filter to the image[20].

We begin as follows $\psi(x)$. It is a step function taking values 1 and -1 on [0,0.5) and [.5,1) respectively as follows:

$$\psi_{j,k}(x) = const \cdot \psi(2^j x - k)$$

define a an orthogonal basis in $L^2(R)$, the space of square integrable functions. This means that any element in $L^2(R)$ may be represented as a linear combination (possibly infinite of these basis functions

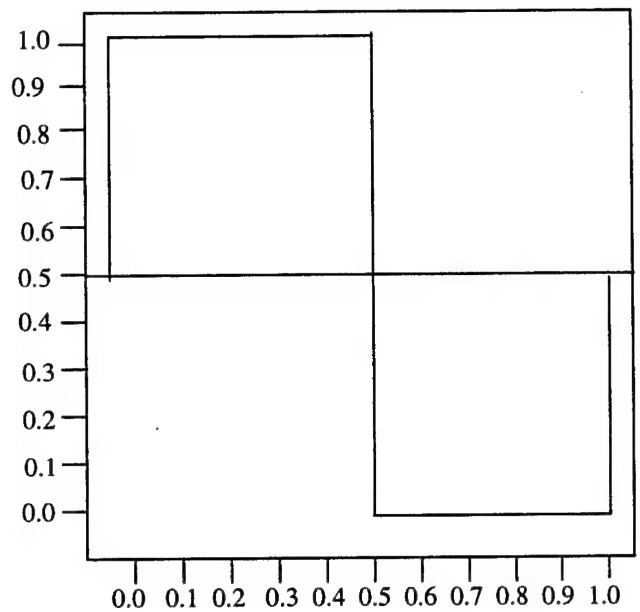


Figure 3-2 Haar Wavelet

A general approach to using the Haar Wavelet to decompose a signal is as follows developed by Stephan Mallat. We start with the space L^2 of square integrable functions. The MRA is an increasing sequence of close subspace $\{V_j\}_{j \in Z}$ which approximates $L^2(R)$.

3.2.1 Basis Derivation

Everything starts with a clever choice of the scaling function ϕ. Except for the Haar wavelet basis for which ϕ is the characteristic function of the interval $[0,1)$, the scaling function is

chosen to satisfy some continuity, smoothness and tail requirements. But, most importantly, the family $\{\phi(x-k), k \in Z\}$ forms an orthonormal basis for the reference space V_0. The following relation describes the analysis:

$$\cdots \subset V_{-1} \subset V_0 \subset V_1 \cdots \tag{3.2.1}$$

The spaces V_j are nested. The space $L^2(R)$ is a closure of the union of all V_j. In other words, $\cup_{j \in Z} V_j$ is dense in $L^2(R)$. The intersection of all V_j is empty.

$$f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}, j \in Z \tag{3.2.2}$$

The spaces V_j and V_{j+1} are "similar". If the space V_j is spanned by $\phi_{jk}(x), k \in Z$ then the space V_{j+1} is spanned by $\phi_{j+1,k}(x), k \in Z$. The space V_{j+1} is generated by the functions $\phi_{j+1,k}(x) = \sqrt{2}\phi_{jk}(2x)$.

We now explain how the wavelets enter the picture. Because $V_0 \subset V_1$, any function in V_0 can be written as a linear combination of the basis functions $\sqrt{2}\phi(2x-k)$ from V_1 in particular:

$$\phi(x) = \sum_k h(k)\sqrt{2}\phi(2x-k) \tag{3.2.3}$$

Coefficients h(k) are defined as $\langle \phi(x), \sqrt{2}\phi(2x-k) \rangle$. Consider now the orthogonal complement W_j of V_j to V_{j+1} (i.e. $V_{j+1} = V_j \oplus W_j$). Define

$$\psi(x) = \sqrt{2}\sum_k (-1)^k h(-k+1)\phi(2x-k) \tag{3.2.4}$$

It can be shown that $\{\sqrt{2}\sum_k \psi(2x-k), k \in Z\}$ is an orthonormal basis for W_1. Again the similarity property of MRA gives that $\{2^{j/2}\psi(2x-k), k \in Z\}$ is a basis for W_j. Since

$\cup_{j \in Z} V_j = \cup_{j \in Z} W_j$ is dense in $L^2(R)$, the family $\psi_{jk}(x) = 2^{j/2}\psi(2x-k), k \in Z$ is a basis for $L^2(R)$.

For a given function $f \in L^2(R)$ one can find N such that $f \in V_N$ approximates f up to some up to preassigned precision (in terms of L_2 closeness). If $g_i \in W_i$ and $f_i \in V_i$, then

$$f_N = f_{N-1} + g_{N-1} + \sum_{i=1}^{M} g_{N-M} + f_{N-M} \tag{3.2.5}$$

The above equation is the wavelet decomposition of f.

3.2.2. QMF structure

The filter bank is the means by which a wavelet is implemented to decompose a signal or image into different frequency bands. There are many filter bank implementations. The quadrature mirror filter (QMF) is here is presented as an example as one of the most basic filter bank implementations and is routinely used for image decomposition [72].

Repeating the above description in[72,73]terms of signal processing we recall that

$$\phi(x) = \sum_{k \in Z} h(k)\sqrt{2}\phi(2x-k) \tag{3.2.6}$$

and

$$\psi(x) = \sum_{k \in Z} g(k)\sqrt{2}\phi(2x-k) \tag{3.2.7}$$

The l^2 sequences $\{h(k), k \in Z\}$ and $\{g(k), k \in Z\}$ are quadrature mirror filters in the terminology of signal analysis. The connection between hand g is given by:

$$g(n) = (-1)^n h(1-n) \tag{3.2.8}$$

The sequence h(k) is the low pass filter while g(k) is the high pass filter. The following properties of h(n), g(n) can be proven by using Fourier transforms and orthogonality:

$\sum h(k) = \sqrt{2}, \sum g(k) = 0$ As a result this filter structure k is shown in Figure 3-3

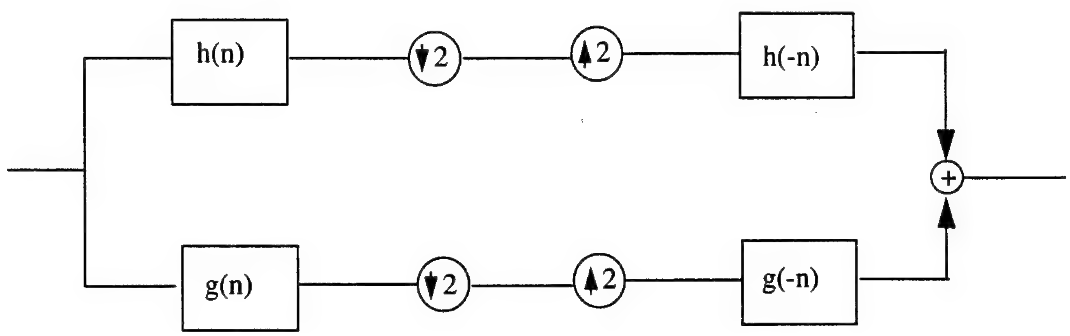


Figure 3-3 Biorthogonal Filter Bank Structure

The most compact way to describe Mallat's multi-resolution analysis MRAand determine the wavelet coefficients is the operator representation filters. Themallatmultiresolutionfrequency structure is shown graphically below in Figure 3-4

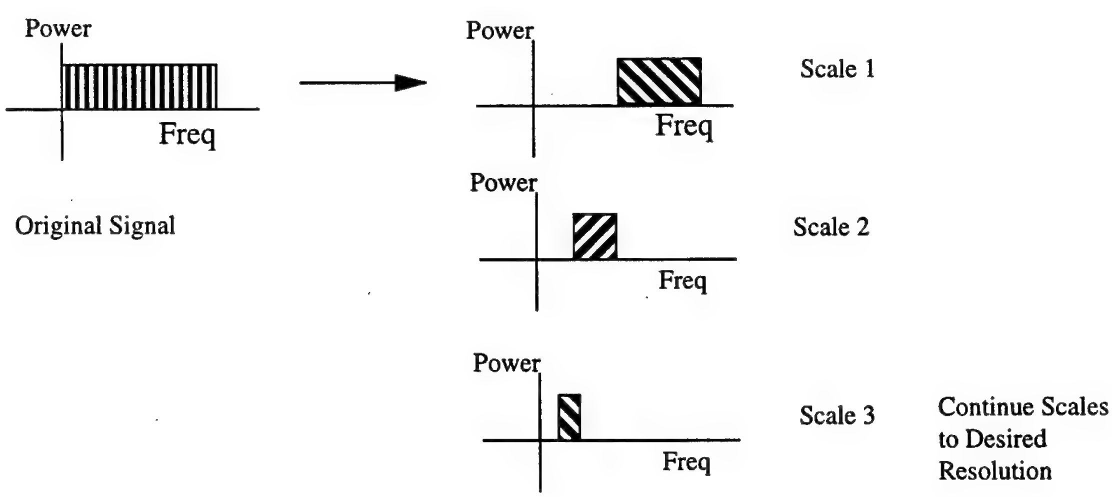


Figure 3-4 Mallat Wavelet Frequency Structure

It is useful to note that this frequency structure of the MRA approach is identical to the Fischer's fractal frequency decomposition shown if Figure 2-4. This makes sense since at each stage of

domain to range mapping on a quad-tree partition essentially applies a Haar H filter to the domain image and then maps this information to the next higher frequency band of range blocks. [20]

For a sequence $a = \{a_n\}$ the operators H and G are defined by the following coordinate-wise relations:

$$(Ha)_k = \sum{}_n h(n-2k)a_n \tag{3.2.9a}$$

$$(Ga)_k = \sum{}_n g(n-2k)a_n \tag{3.2.9b}$$

The operators H and G correspond to one step in the wavelet decomposition. The only difference is that the above definitions do not include the $\sqrt{2}$ factor.

Denote the original signal by $c^{(n)}$. If the signal is of length 2^n, then $c^{(n)}$ can be represented by the function $f(x) = \sum c_k^{(n)} \phi_{nk}, f \in V_n$. At each stage of the wavelet transformation we move to a coarser approximation $c^{(j-1)}$ by $c^{(j-1)} = Hc^{(j)}$ and $d^{(j-1)} = Gc^{(j)}$. Here, $d^{(j-1)}$ is the detail lost by approximating is the detail lost by approximating $c^{(j-1)}$. The discrete wavelet transformation of a sequence of length 2^n (notice that the sequence $c^{(j-1)}$ has half the length of $c^{(j)}$):

$$(d^{(n-1)}, d^{(n-2)}, \ldots\ldots, d^{(1)}, d^{(0)}, c^{(0)}) \tag{3.2.10}$$

Thus the discrete wavelet transformation can be summarized as a single line:

$$y \rightarrow (Gy, GHy, GH^2y, \ldots\ldots GH^{n-1}y, GH^n y) \tag{3.2.11}$$

The reconstruction formula is also simple in terms of H and G; we first define a joint operators H^* and G^* as follows:

$$(H^* a)_n = \sum{}_k h(n-2k) \tag{3.2.12a}$$

$$(G^{*}a)_n = \sum{}_k h(n-2k) \qquad (3.2.12b)$$

Recursive application leads to:

$$(Gy, GHy, GH^2y, ...GH^{n-1}y, GH^n y) \rightarrow y = \sum_{j=0}^{n-1} (H^*)^j G^* d^{(j)} + (H^*)^n c^{(0)} \qquad (3.2.13)$$

The above equations (generating equations) used with the Haar wavelet are as follows:

$$\phi(x) = \phi(2x) + \phi(2x-1) = \frac{1}{\sqrt{2}}\sqrt{2}\phi(2x) + \frac{1}{\sqrt{2}}\sqrt{2}\phi(2x-1) \qquad (3.2.14)$$

$$\psi(x) = \phi(2x) + \phi(2x-1) = \frac{1}{\sqrt{2}}\sqrt{2}\phi(2x) + \frac{1}{\sqrt{2}}\sqrt{2}\phi(2x-1) \qquad (3.2.15)$$

The filter coefficients in the above equations are:

$$h(0) = h(1) = \frac{1}{\sqrt{2}} \quad g(0) = -g(1) = \frac{1}{\sqrt{2}} \qquad (3.2.16)$$

To get the wavelet coefficients we multiply the components of $d^{(j)}$, j = 0,1,2 and $c^{(0)}$ with the factor $2^{-N/2}$. Simply:

$$d_{jk} = 2^{-N/2} d_k^{(j)}, 0 \le j \le N(=3) \qquad (3.2.17)$$

It is interesting that the Haar wavelet case $2^{-3/2}c_0 = c_{00} = 0.5$ is the mean of the sample y.

The entire frequency bank decomposition is shown below in Figure 3-5

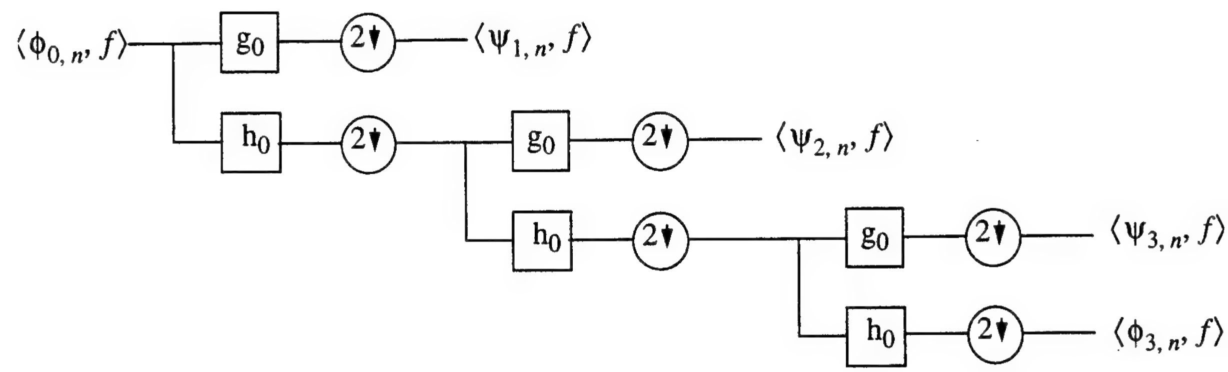


Figure 3-5 Multiresolution QMF Frequency Decomposition

Thus we have a direct description of how the fractal image compression process divides itself in terms of frequency representation by means of the QMF Haar frequency decomposition.

3.2.2 2-D representation

The QMF two dimensional decomposition [73] is shown below. The give image is decomposed in each direction by a high and low pass filter and downsampled in those directions. Next the wavelet is decomposed in the other direction in the same manner for a resulting two

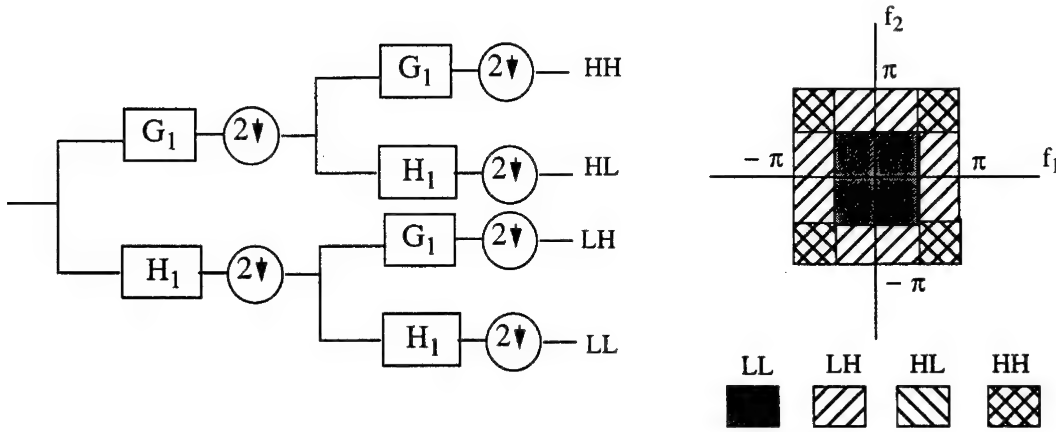


Figure 3-6 Two Dimensional Multiresolution QMF

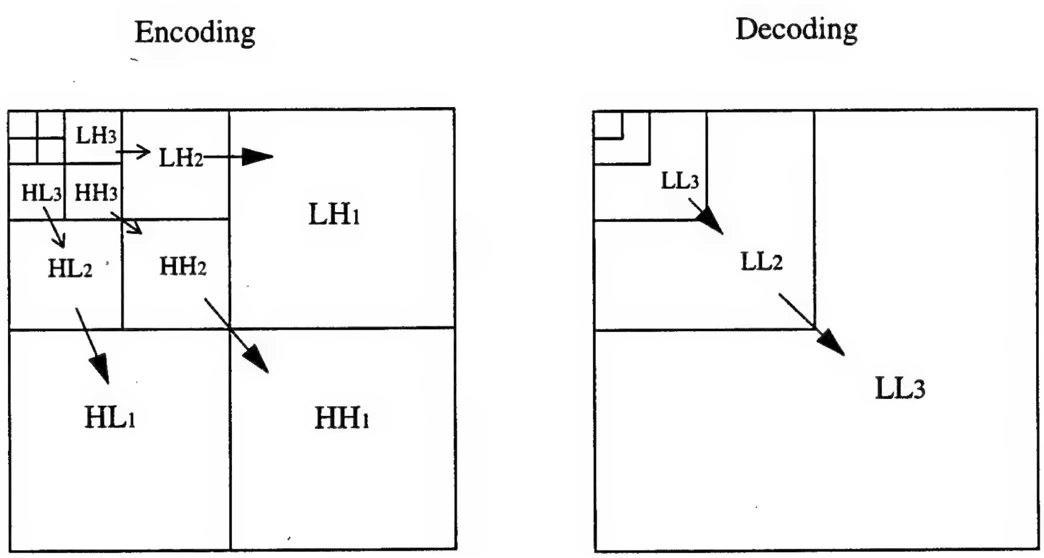
dimensional frequency spectra as is shown above. While this method is the standard image decomposition used in much of wavelet image processing the Haar basis set itself has many drawbacks in the localization of its frequency spectra and the subsampling process decreases the inherent resolution of the image decomposition[20].

3.3 Existing Fractal-Wavelet Analog Techniques

Recently a number of researchers including Geoff Davis noted that the Haar wavelet showed a direct analog in the fractal compression process[20, 57]. As a result the method of the Haar transform has recently been adapted to improve the fractal coding processs. In addition another wavelet compression technique has emerged which like the fractal uses cross scale redundancy to achiev compression This method Zero Tree Method known as the zero tree was developed by Shapiro [59] and uses a wavelet decomposition similar to the QMF Harr decomposition described earlier. These techniques also give insight on how to design a better compression method using fractal and wavelet principles.

3.3.1 Davis SQS Technique

Geoff Davis[20, 57] and others noted that the wavelet transform could mimic the block mapping process of the fractal transform and developed the first hybrid transform. First Davis realized that the averaging process in range to domain block mapping was the low-pass wavelet of the Haar transform. He also noted that the range to domain mapping process at each scale could be made significantly faster by matching the high pass bands to each other to approximate the low pass to high pass reconstruction . This frequency division makes sense since the high pass must be



3-7 Frequency Analysis of Davis' Technique

approximated upon fractal reconstruction anyway, so the frequency division process of the wavelet transform simply organizes this process. Davis' method also shows outstanding frequency compression as well. Unfortunately, the complexity of the compression scheme in this case is high since there are multiple high pass bands to map from high pass to low pass on encoding. Also

Davis does not take into account the implicit relationships between shape and texture In addition this technique does not preserve shapes in the compressed domain.

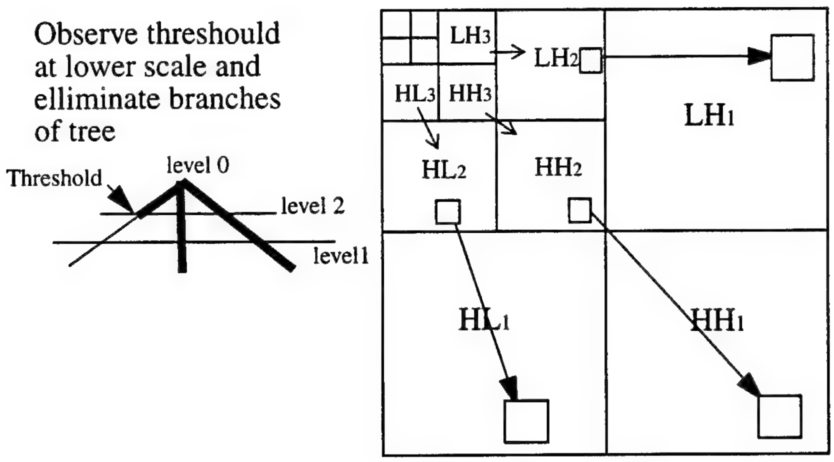
3.3.1.1 Band structure mapping

Figure 3-7 shows how the local IFS of the hybrid techniques maps Domain to Range Blocks from higher to lower scales. This process is the same as the multifractal except the mapping changes from scale to scale since the wavelet fractal process uses a local rather than global iterated function system. The reconstruction process we will first develop then uses the attractor model as in the multifractal case to reconstruct. Thus we have a linear synthesis and nonlinear piecewise reconstruction.

As is also shown in Figure 3-7 the reconstructive mapping approximates higher frequency information with information from lower scales. This process thus is a predictive coding technique from low to high frequency.

3.3.2 Shapiro Zero Tree

The Shapiro Zero Tree [59] algorithm has been called “the state of the art in image coders”. Essentially zero-tree uses the fact that power across frequency scales tends to naturally decay as one increases in frequency or decreases in scale. This fact allows a person attempting image coding to eliminate information in image at a particular frequency scale and all scales below it by determining the local energy in a block in one of the high pass images. If the energy is below a certain threshold the coefficients at that scale and all scales below it are not coded.



3-8 Schapiro's Zero Tree

3.4 Biorthogonal Spline

The Haar in the applications above is a useful structure because of its simplicity in implementation. However, for content based image compression we look to a different wavelet structure to directly reveal the natural shape and texture composition of an image. This structure is the biorthogonal spline wavelet.

Unlike the orthogonal structure, the biorthogonal structure reconstructing wavelet is not the complex conjugate of the analyzing wavelet but a totally separate function. Such biorthogonal structures are useful in maintaining symmetry in the wavelet filters. Such symmetry is necessary later for such applications as the edge detection routines used in Mallat's [45] decomposition shown later.

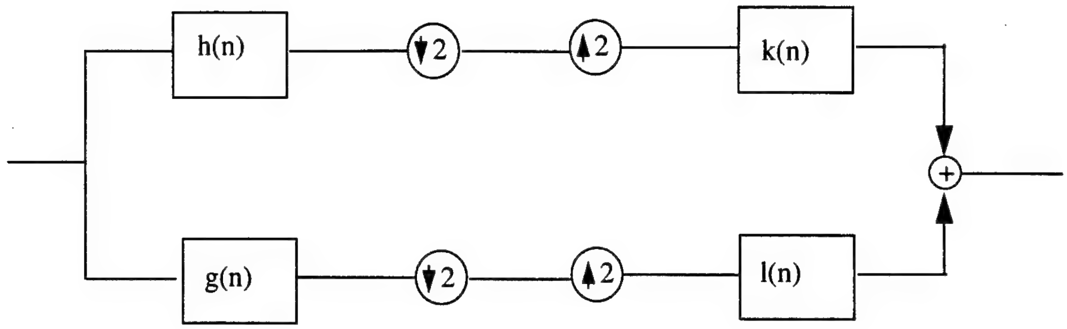


Figure 3-9 Biorthogonal Filter Bank Structure

3.4.1 Gaussian Derivative Basis

A particular implementation of the biorthogonal spline uses the Gaussian derivative basis set. As we shall show later the Gaussian derivative has many desirable properties useful in signal and image analysis including an ability to accurately reveal boundaries and edges in signals and imagery minimal artifact production in image reproduction. These properties will be extremely useful in designing a new fractal wavelet method of image analysis and synthesis in Chapters 4 and 5.

The Mallat transform $\psi(x)$ is constructed in a very different manner from the QMF Haar. Start with a gaussian function and take its derivative as shown in the following figures:

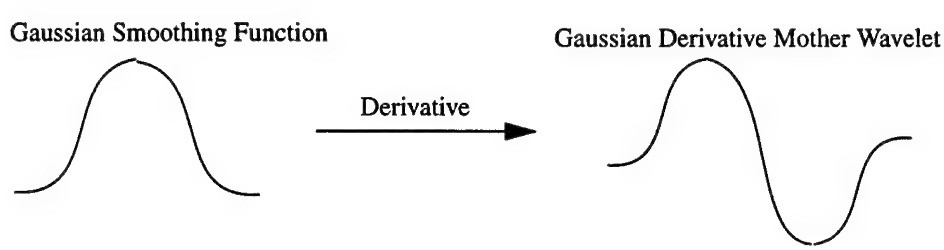


Figure 3-10 First Derivative Gaussian Smoothing Function

The result is a an extremely effective means of detecting singularities(edges) in a signal. If we represent our gaussian $g_{\sigma}(x)$

$$g_{\sigma}(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{x^2}{2\sigma^2}\right) \tag{3.4.1}$$

then our smoothing function at scale s_0 is :$\theta_{s_0}(x)$ where we assume:

$$\theta_{2^j} * g_{\sigma}(x) \approx \theta_{s_0}(x) \tag{3.4.2}$$

Our defining wavelet is thus

$$\psi(x) = \frac{\partial}{\partial x}\theta(x) \tag{3.4.3}$$

For each scale we have

$$\psi_s(x) = \frac{1}{s^2}\psi\left(\frac{x}{s}\right) \tag{3.4.4}$$

Our 1-D wavelet operator is thus:

$$W_s^1 f(x) = f \otimes \psi_s(x) \tag{3.4.5}$$

3.4.2 Wavelet Implementation

We impose that the Fourier transform [73] of the smoothing function $\phi(x)$ defined by

$$|\hat{\phi}\langle w\rangle|^2 = \sum_{j=1}^{+\infty} \hat{\psi}(2^j w)\hat{\chi}(2^j w) \tag{3.4.6}$$

can be written as an infinite product

$$\hat{\phi}(w) = e^{-iw\omega} \prod_{p=1}^{+\infty} H(2^{-p\omega}) \tag{3.4.7}$$

where $H(\omega)$ is a 2π periodic differentiable function such that

$$|H(\omega)|^2 + |H(\omega+\pi)|^2 \le 1 \text{ and } |H(0)| = 1. \tag{3.4.8}$$

One can prove that the conditions are sufficient so that the above equations define a smoothing function $\phi(x)$, which is in $\mathbf{L}^2(\mathbf{R})$. The parameter w is a sampling shift. It its adjusted in order that $\phi(x)$ is symmetrical with respect to 0. The above equations imply that

$$\hat{\phi}(2\omega) = e^{-iw\omega} H(\omega)\hat{\phi}(\omega) \tag{3.4.9}$$

We define a wavelet $\psi(x)$ whose Fourier transform $\tilde{\psi}(x)$ is given by

$$\hat{\psi}(2\omega) = e^{-iw\omega} G(\omega)\hat{\phi}(\omega) \qquad (3.4.10)$$

where $G(\omega)$ is a 2π periodic function. Let us impose that the Fourier transform of the reconstructing wavelet $\tilde{\chi}(x)$ can be written

$$\hat{\chi}(2\omega) = e^{iw\omega} K(\omega)\hat{\phi}(\omega) \qquad (3.4.11)$$

From the above equations we can derive

$$G(\omega)K(\omega) + |H(\omega)|^2 = 1 \qquad (3.4.12)$$

and for the two dimensional transform

$$L(\omega) = \frac{1 + |H(\omega)|^2}{2} \qquad (3.4.13)$$

We want a wavelet $\psi(x)$ equal to the first order derivative of a smoothing function $\theta(x)$. This implies that $\tilde{\psi}(x)$ must have a zero of order 1 at $\omega = 0$. We choose $H(\omega)$ in order that a wavelet $\psi(x)$, which is antisymetrical, as regular as possible , and has a small compact support. A family of 2π periodic functions that satisfy these constraints is given by.

$$H(\omega) = e^{i\omega/2}(\cos(\omega/2))^{2n+1} \qquad (3.4.14)$$

$$G(\omega) = 4ie^{i\omega/2}\sin(\omega/2) \qquad (3.4.15)$$

We can also derive

$$\hat{\phi}(\omega) = \left(\frac{\sin(\omega/2)}{\omega/2}\right)^{2n+1} \tag{3.4.16}$$

and

$$\hat{\psi}(\omega) = i\omega\left(\frac{\sin(\omega/4)}{\omega/4}\right)^{2n+2} \tag{3.4.17}$$

The Fourier transform of $\hat{\theta}(x)$ of the primitive is therefore.

$$\hat{\theta}(\omega) = \left(\frac{\sin(\omega/4)}{\omega/4}\right)^{2n+2} \tag{3.4.18}$$

The overall filter bank structure using the above wavelet basis is shown in Figure 3-9 and is again based on Mallat's multiresolution analysis.

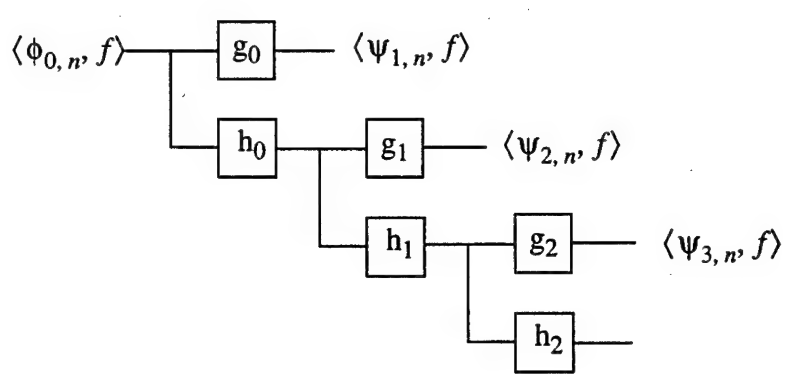


Figure 3-11 Multiresolution Biorthogonal Frequency Decomposition

3.4.3 Fast Implementation

A fast iterative method for the above wavelet decomposition is described below. This fast method is pivotal in ultimately constructing a fast multiresolution decomposition as well as fast image reconstruction using our new method of Chapter 5.

We suppose that the wavelet $\psi(x)$ is [44] characterized by the three discrete filters H, G, and K and L. We denote H_p, G_p, and K_p and L_p. the discrete filter obtained by putting 2^p-1 zeros between consecutive coefficients of the filters. We also denote by D the Dirac filter whose impulse response is equal to 1 at 0 and 0 otherwise. We denote by A*(H,L) the separable convolution of the rows and columns, respectively of the image A with 1-D filters H and L.

j = 0

while (j < J)

$$W^{1,d}_{2^{j+1}} f = \frac{1}{\lambda_j} S^d_{2^j} f * (G_j, D) \tag{3.4.19a}$$

$$W^{2,d}_{2^{j+1}} f = \frac{1}{\lambda_j} S^d_{2^j} f * (D, G_j) \tag{3.4.19b}$$

$$S^d_{2^{j+1}} f = \frac{1}{\lambda_j} S^d_{2^j} f * (H_j, H_j) \tag{3.4.19c}$$

j = j + 1

endwhile

A fast iterative reconstruction process is also described below

j = J

while (j > 0)

$$S_{2^{j+1}}^{d} f = \lambda_j \cdot W_{2^j}^{1,d} f*(K_{j+1}, L_{j+1}) + \lambda_j \cdot W_{2^j}^{2,d} f*(L_{j+1}, K_{j+1}) +$$

$$\frac{1}{\lambda_j} S_{2^j}^{d} f*(\tilde{H}_{j+1}, \tilde{H}_{j+1}) \qquad (3.4.20)$$

j = j - 1

endwhile

Specific coefficients for the implementation of these filters are shown in the following table:

Finite Implulse Response of the Filters H, G, K and L That Correspond to The Quadratic Spline Wavelet

n	H	G	K	L
-3			0.0078125	0.0078125
-2			0.054685	0.046875
-1	0.125		0.171875	0.1171875
0	0.375	- 2.0	- 0.171875	0.65625
1	0.375	2.0	- 0.054685	0.1171875
2	0.125		- 0.0078125	0.046875
3				0.0078125

Table 3-1

3.4.4 2-D representation

In contrast to the Haar QMF orthogonal 2-D decomposition shown in section 3.2 Mallat uses a more elegant technique for two dimesional decomposition which lends itself to image compression and is the basis of the Hybrid Fractal Wavelet Method.

The algorithm consists of first preprocessing an image with a multiscale wavelet decomposition as described by Mallat [45] . We chose the number of scales to decompose based on the resolution of the image we wish to reproduce and the amount of edge information we wish to save (Fig. 8.) Wavelet encoding in 2 dimensions is generically described in equations 3.4.21 through 3.4.24.

$$\psi^1(x,y) = \frac{\partial}{\partial x}\theta(x,y) \qquad \psi^2(x,y) = \frac{\partial}{\partial y}\theta(x,y) \tag{3.4.21}$$

$$\psi_s^1(x,y) = \frac{1}{s^2}\psi^1\left(\frac{x}{s},\frac{y}{s}\right) \qquad \psi_s^2(x,y) = \frac{1}{s^2}\psi^2\left(\frac{x}{s},\frac{y}{s}\right) \tag{3.4.22}$$

$$W_s^1 f(x,y) = f \otimes \psi_s^1(x,y) \qquad W_s^2 f(x,y) = f \otimes \psi_s^2(x,y \tag{3.4.23}$$

$$Wf = (W_s^1 f(x,y), W_s^2 f(x,y))_{j \in Z} \tag{3.4.24}$$

Ψ^1 and $\Psi^2$are the gradients of the smoothing function θ in the x and the y directions respectively. $\Psi_s{}^1$ and $\Psi_s{}^2$are the gradients at each scale s which are usually power of 2 in the spatial x and y directions. $W_s{}^1$ and $W_s{}^2$ are the wavelet transform functions in the x and y directions.

3.4.5 Modulus Maxima

After computing the wavelet transform [44] we compute modulus and gradient angle images as described by equations 3.3.25 and 3.3.26.

$$M_s f(x,y) = \sqrt{\left|W_s^1 f(x,y)\right|^2 + \left|W_s^2 f(x,y)\right|^2} \tag{3.4.25}$$

$$A_s f(x,y) = \arg\left(W_s^1 f(x,y) + i W_s^2 f(x,y)\right) \tag{3.4.26}$$

The difference between Mallat in his two dimensional biorthogonal transform and the ordinary orthogonal Haar two dimensional transform is that he does not subsample his images as in the 1 dimesional case and that he applies only one filter in the X and Y directions to compute a polar representation for modulus maxima for each of the high pass bands. Thus, Mallat has only two high pass bands which can be represented in terms of x and y or polar representation.

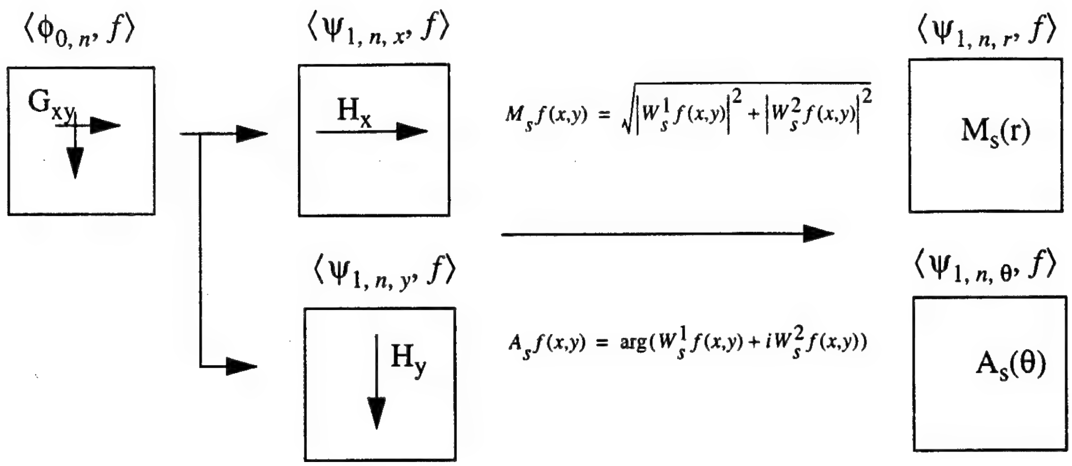
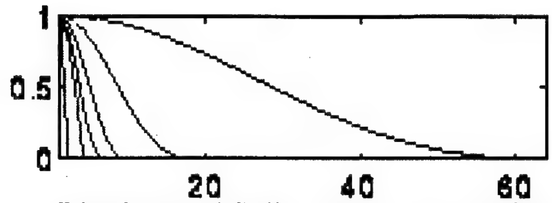
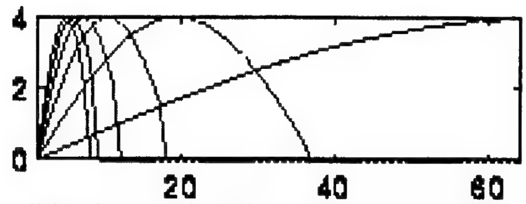


Figure 3-12 Modulus Maxima

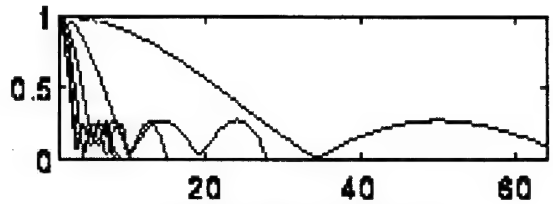
It is also interesting to note as shown in Figure 3-13 that the biorthogonal spline has a much more accurate frequency response than the Haar basis set. This fact is extremely important in the wavelet reconstruction process since the Haar basis contains many high frequency artifacts in reproduced images due to its sharp spatial domain cutoff. The Gaussian derivative spline has a much smoother spatial cutoff and thus much less tendency to create artifacts in imagery. as is shown in the frequency plots below:



Biorthogonal Spline Frequency Profile
Low Pass H Filters Scale 1-6 1-D



Biorthogonal Spline Frequency Profile
G Filters Scales 1-6 1-D



Haar Frequency Profile
Low Pass H Filters Scales Scales 1-6 1-D

Figure 3-13 Mallat's Gaussian Derivative vs. Harr

3.4.6 Mallat Technique

In a technique developed by Stephan Mallat and Sifen Zhong,[44] they use the modulus maxima edge information from the 2 dimensional biorthogonal spline wavelet described above to reconstruct the original image. Thus modulus maxima lines are chain coded and only those with intensity above a certain threshold are store as in the Zero Tree example.

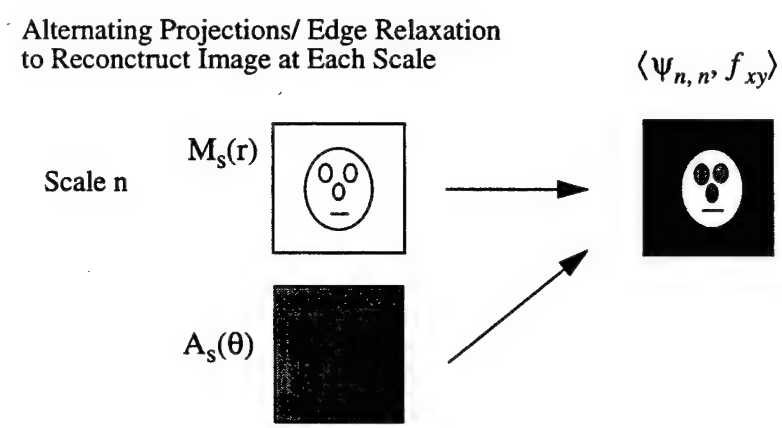


Figure 3-14 Alternating Projections Reconstruction

The synthesis process for reconstructions involves a complicated alternating projection process which is similar to nonlinear edge relaxation. While the edge information has great use in analyzing shapes in imagery the reconstruction process can be somewhat time consuming and also has a tendency to remove image texture since it is essentially a nonlinear interpolation process.

Chapter 4

Multifractal Model

Contents

4.1 Multifractal 1-D Signal Decomposition/Reconstruction

In order to define our compression technique in terms of the images natural shape and texture we turn to an approach to signal decomposition known as the multifractal method. Recently the multifractal model has emerged as a defining connection between wavelets and fractals. The multifractal in most implementations uses Mallat's MRA with biorthogonal spline wavelet decomposition to analyze a signal and a global iterated function system invariant measure algorithm to reconstruct the signal. We will first derive how the multifractal connects both wavelet and fractal models. We will then show how parameters from this technique may be used to describe the natural fractal shape and texture of an object. Using this connection we will ultimately develop the wavelet-fractal compression method.

To understand the multifractal concept we first give an example of an invariant measure one dimensional iterated function system. This will serve as the basis for synthesis of any arbitrary signal. This method is a mathematical example of the method described in section 2.2.1.1.

4.1.1 An Invariant Measure Example

The contractions $w_1, w_2, \ldots\ldots, w_N$ and the [53] probabilities $p_1, p_2, \ldots\ldots, p_N$ determine how frequently a certain pixel P_{ij} will be hit by the chaos game. The average fraction

$$\lim_{k \to \infty} \frac{h(z_1, \ldots\ldots, z_k; P_{ij})}{k} = R_{ij} \tag{4.1.1}$$

if the result of a particular measure μ which has the attractor A_∞, the attractor of the IFS as its support (i.e. $\mu(A_\infty) = 1$). In other words,

$$\mu(P_{ij}) = R_{ij} \tag{4.1.2}$$

This measure μ is Borel measure and is invariant under the Markov operator $M(v)$ which is defined as mentioned in chapter 1 with slightly different notation. Let X be a large square in the plane which contains A_∞, the attractor of the IFS, and v a (Borel) measure on X. Then this operator is defined by

$$M(\nu) = p_1 \nu w_1^{-1} + p_2 \nu w_2^{-1} + \cdots + p_N \nu w_N^{-1} \tag{4.1.3}$$

In other words, $M(\nu)$ defines a new normalized Borel measure on X. We evaluate this measure for a given subset B in the following way: first we take the preimages $w_i^{-1}(B)$ with respect to X, then evaluate v on that and finally we multiply the probabilities p_i and add up the results. Here is an example. Let

$$w_1(x) = (1/2)x, \qquad p_1 = 1/3 \tag{4.1.4}$$

$$w_2(x) = 1/2x + 1/2, \qquad p_2 = 2/3 \tag{4.1.5}$$

This is an IFS which has the unit interval as its attractor $A_\infty = [0, 1]$. Now assume that we start with a measure given by the density:

$$h_0(x) = \begin{cases} 1 & \text{if } x \in [0, 1] \\ 0 & \text{otherwise} \end{cases} \tag{4.1.6}$$

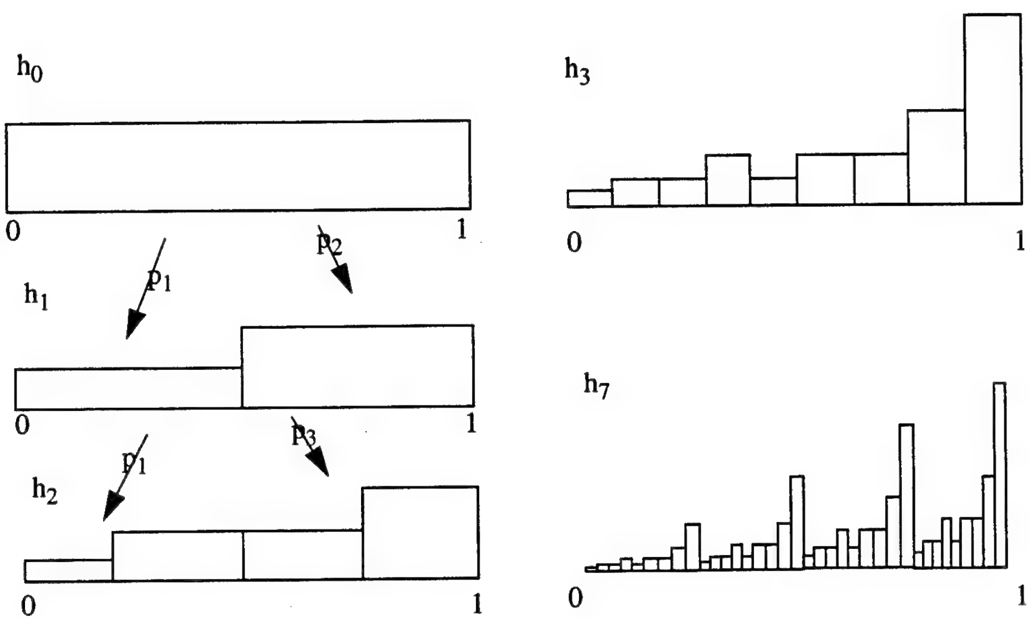


Figure 4-1 Invariant Measure Example

i.e., the initial measure is $\nu_0 = \int_A h_0(x)dx$. For a subset $A \subset [0, 1/2]$ of the left half unit interval we have $w_2^{-1}(A) \subset [-1, 0]$ and $\nu_0(w_2^{-1}(A)) = 0$. Thus $\nu_1(A) = \nu_0(w_2^{-1}(A))$. A corresponding argument holds for the right half interval [1,1/2] while w_2 does the same with the right half interval multiplying the result by p_2. Thus after the first step we obtain the density

$$h_1(x) = \begin{cases} p_1 & \text{if } x \in [0, 1/2] \\ p_2 & \text{if } x \in [1/2, 0] \\ 0 & \text{otherwise} \end{cases} \tag{4.1.7}$$

and $\nu_1(A) = \int_A h_1(x)dx$. We construct a measure $\nu_2 = M(\nu_1)$ along the same lines and obtain the density function h_2 as shown in the above figure. In the limit this process generates a binomial measure that is a self-similar multifractal measure In the next section we will describe a new selection of functional equations incorporating the wavelet from which we will be able to reveal the properties of an invariant measure and synthesize any arbitrary signal.

4.1.2 Self Similar Wavelet Functional Equation

To determine the invariant measure implicit in a given signal we have a method of interpreting its essential temporal or spatial structure. This structure is defined by the singularities or edges of the signal. Thus, to detect singularities [36] in the original signal we can find the points within the signal where most energy is localized using the biorthogonal spline approach with Mallat's MRA. The next task is to find a mapping or set of iterated function system equations that characterizes the original signal.

Note that the mapping proceeds from the lowest frequency band to the next higher band of frequencies. [1] The goal of the multifractal process is to use the wavelet transform to reveal the invariant measure parameters of probability p , translation r, and scale l. If we have a self similar signal *f*(x) which can be approximated the invariant measure parameters as follows:

$$f(x) = plf(ls, (x-r)) \tag{4.1.8}$$

We can show that by taking the wavelet transform of 4.1.8 that these these parameters naturally

fall out of the wavelet transform as follows:

$$Wf(s, x) = plWf(ls, l(x-r)) \tag{4.1.9}$$

Or more importantly taking the wavelet maxima

$$Mf(s, x) = plMf(ls, l(x-r)) \tag{4.1.10}$$

By using the wavelet maxima singularity spectra shown in Fig 4-2c this process can be greatly simplified since we have separated the original signal into frequency bands and then reduced the information within these bands to an elementary set of data, namely the maxima themselves. We can characterize the original equations that resulted in figure 4-2c by a histogram voting procedure that records the following quantities l, p and r directly from the maxima representation. l as indicated in 4.1.10 characterizes the contractivity factor or spacing between separate wavelet scales

$$. \log l = \log\left(\frac{s_2}{s_1}\right) \tag{4.1.11}$$

p is probability of the invariant measure and characterizes the decay of energy across frequency bands.

$$\log p = \log\left(\frac{Wf(s_1, x_1)}{lWf(s_2, x_2)}\right) \tag{4.1.12}$$

r is the geometrical displacement or translation between scales

$$r = x_1 - x_2 \frac{s_1}{s_2} \tag{4.1.13}$$

4.1.3 Analysis with wavelet -> Reconstruction with fractal

The process of multifractal analysis and construction amounts to decomposition with the wavelet and invariant measure or ifs reconstruction. To detect singularities we use a wavelet basis function in the time domain which follows the above frequency structure but in the convolution process detects singularities or "edges" in imagery. Such a basis set might be the derivative of a Gaussian function [7] This process is demonstrated in the following three figures.Figure 4-2a shows the Devils Staircase in 1 Dimension created with the functional equations of section 4.1.2 characterized by the multifractal parameters: l_1=2, l_2=4, p_1=0.66, p_2=0.33, r_1=-.25, r_2=0.37

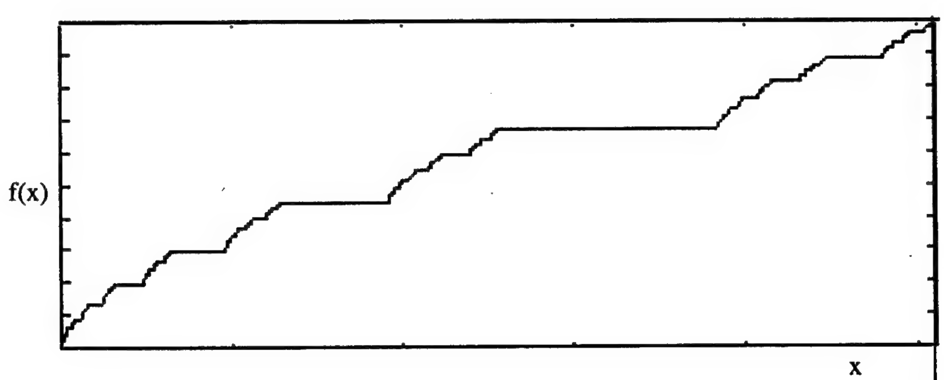
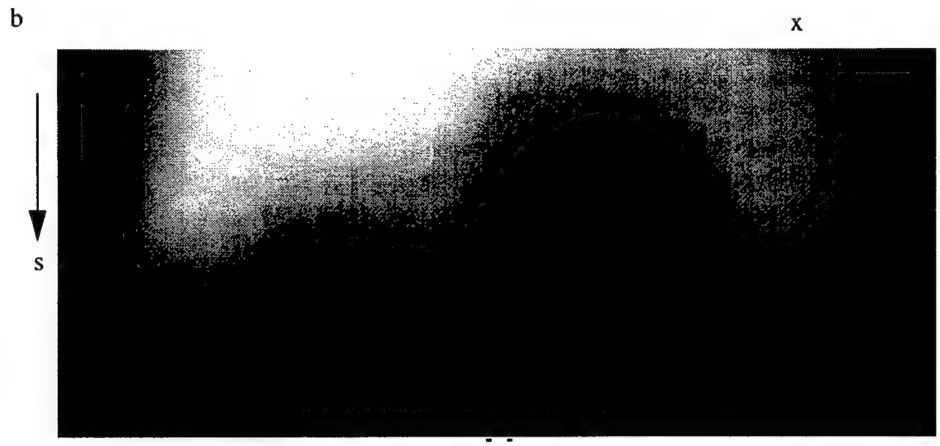
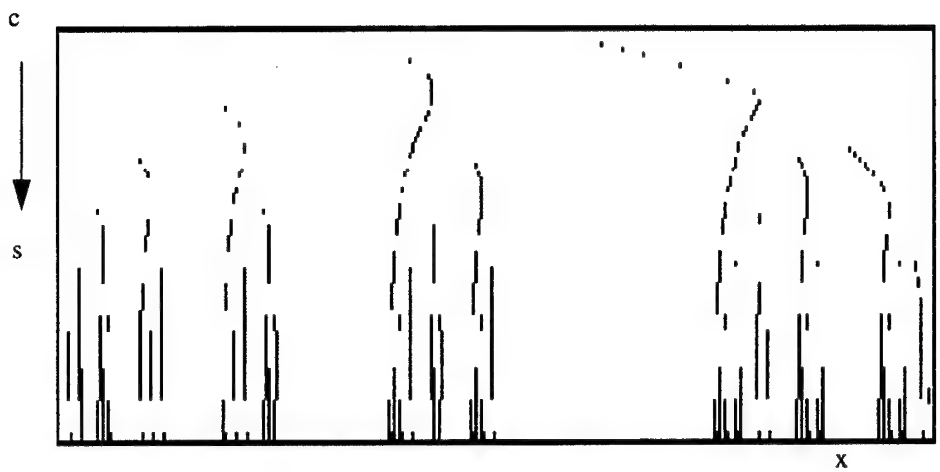


Figure 4-2a Devils Staircase Original Function



Wavelet 1-D Modulus Maxima



Wavelet 1-D Modulus Maxima

Figure 4-2a-c Multifractal Decomposition 1-D

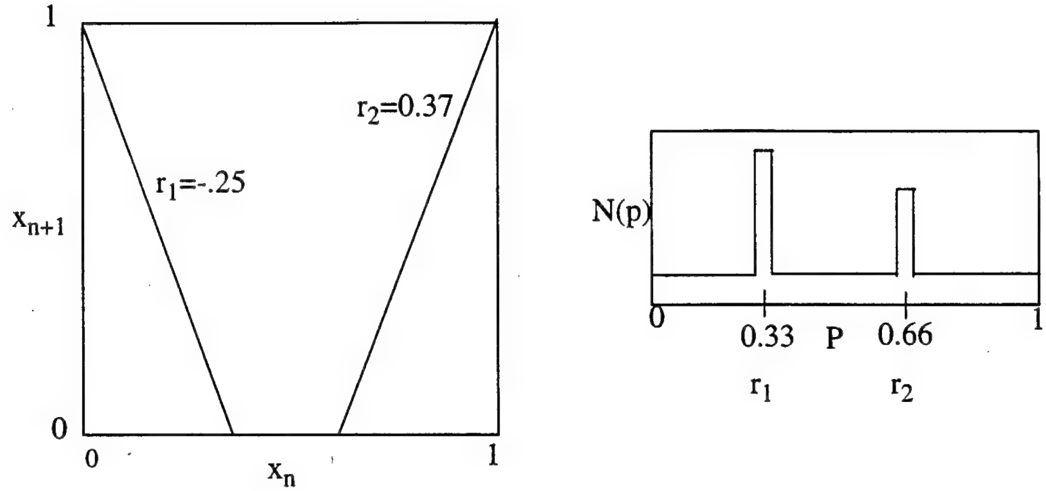


Figure 4-3 Multifractal Parameter Analysis

Figures 4-3 shows the ratios of the spacing between maxima lines in figure 4-2c of the devils staircase followed by histograms of the number of maxima lines which fall into each category. As we can see from histograms of the modulus maxima graphs we are able to derive the exact parameters of the functional equations that created the Devils Staircase of Figure 4-2a. Thus we have a means of breaking apart a given one dimensional signal into its elementary fractal structure and then reconstructing it using iterated function system equations. Needless to say, this method is an extremely effective compression method since it uses the innate mathematical structure of a fractal to compress It also shows the implicit connection between wavelets and fractals. Unfortunately this method is much more challenging in two dimensions. We will now discuss a methodology for approaching this "inverse fractal" problem in two dimensions.

4.1.4 Two Dimensional Global IFS

In two dimensions the iterated function system [11,36] equations have both x and y components for the reconstruction. Such functional equations take a somewhat different form since x and y axis can be coupled. Simplify the process we do not couple the axis and we thus get the two following equations as in the previous section and also assuming that we can approximate our signals f(x) and f(y) with a summation of versions of itself we have:

$$f(x) = \sum_{i=1}^{n} p_i l_i f(l_i(x - r_i)) \tag{4.1.14}$$

$$f(y) = \sum_{i=1}^{n} p_i m_i f(m_i(y - s_i)) \tag{4.1.15}$$

A set of the above equations applied with probabilities p can be used to generate an arbitrary two dimensional function as was done with the devils staircase in 1 dimension. Now setting l_1=.5, l_2=.5, l_3=.5, m_1=0.5, m_2=0.5, m_3=0.5, r_1=0.0, r_2=0.5, r_3=0.0, s_1=0.0, s_2=0.0, s_2=0.5, p_1=0.33, p_2=0.33, p_3=0.33 we produce the Serpinski triangle shown in Figure 4-4

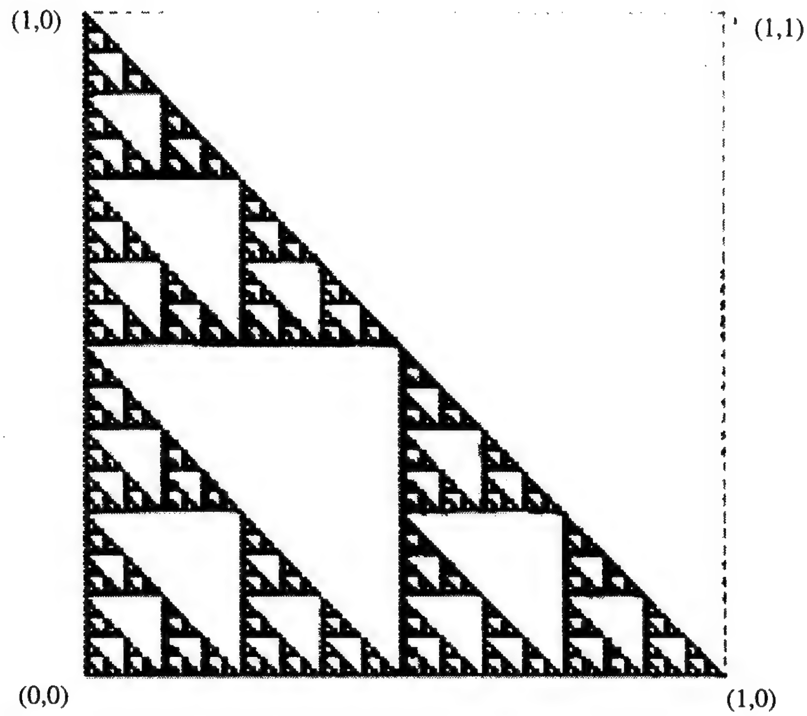


Figure 4-4 Sierpinsky Triangle

Unfortunately the task of finding the two dimensional functional analysis parameters is not as easy as for the 1 dimensional case since the extra dimension adds a geometrical order of magnitude of complexity. We now compare the implementation of multifractal global IFS to local IFS.

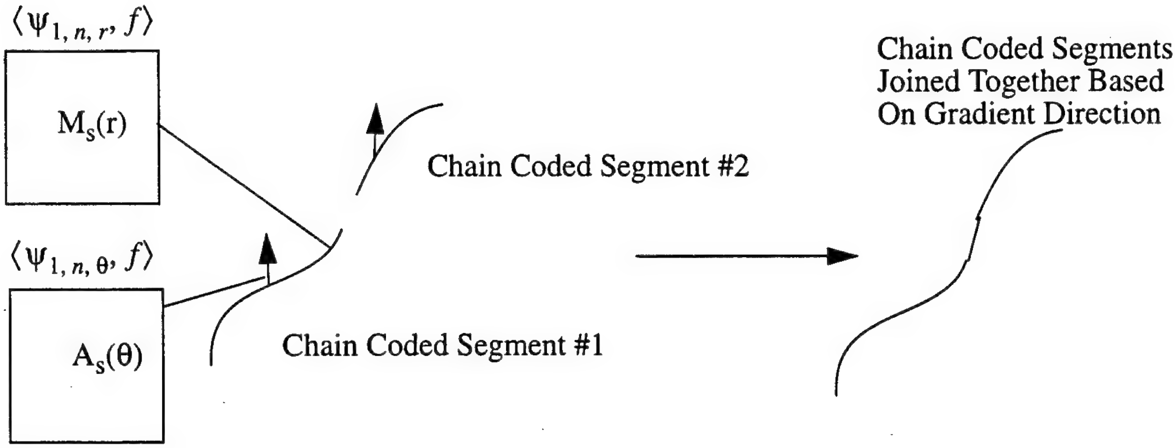
4.2 Two Dimensional Shape and Texture Analysis

Now with an understanding of the multifractal concept we turn to analysis of shape and texture in imagery. First we will examine how the two dimensional singularity spectra reveals the shape of objects in imagery. Then we will apply this singularity spectra to texture and show how the concepts relate to the local IFS model to ultimately improve the image compression process.

4.2.1 2-D Modulus Maxima Reveals Shape

The Mallat basis set is based [14,43,44] on an approximation of the first derivative of a gaussian. Such a filter is shown to have very accurate edge detecting capabilities. In other words the edges detected at each scale are not significantly displaced from where the edge is physically located in an image. Such a property is very useful for image reconstruction as will be shown later. The biorthogonality property of the Mallat Transform significantly helps this process since it allows symmetric basis functions which also produce accurate edges. The Gaussian is also known to produce edges which are consistent and do not represent false edge points. Also the first derivative only produces one edge response per edges unlike higher order derivatives that produce multiple responses.

The zero crossing representation [14,43] produced by the first derivative of the Gaussian is significant because it represents boundaries in 2 dimensional signals between different intensities. These boundaries are revealed by the modulus maxima edges. This modulus maxima edge method is among the most reliable way to detect edges since edges that might otherwise end due to low intensity can be regularized by chain coding and finding continuation of edges based on edge gradient values. Thus Mallat is able to exploit this fact for greater localization of edges across scales. This localization reveals the inherent self-similar structure in an image and our goal in the new wavelet fractal method is to exploit the modulus maxima representation to simplify the fractal compression process in a similar manner to what was achieved in the 1-dimensional multi-fractal case.



4-5 Gradient-Based Chain Coding

4.2.2 Lipschitz Alpha to Characterize Edges

To positively identify edges, Mallat [44] makes use of the multiple frequency scales to see edge patterns over a range wavelet frequency bands . If these edge patterns appear over multiple frequency bands then an edge is positively identified. To characterize the self similarity between scales Mallat uses the Lipschitz criteria as described in equation 4.2.1. If we select some α where $0 < \alpha < 1$ and the function f(x,y) is uniformly Lipschitz over an open set of reals if there exists a constant K such that for all points (x,y) of this open set

$$M_{2^j} f(x, y) \le K(2^j)^{\alpha} \tag{4.2.1}$$

Essentially the α criterion measures the intensity of the wavelet modulus as one progresses to successively higher scales (lower frequencies). If in a two dimensional signal an object has a small exponent, the intensity of the wavelet modulus maxima stays relatively constant over a number of scales and we have essentially a 'hard edge' [7] which stays the same from scale to scale whereas higher α indicate softer edges. Thus in 2-D low α can be used to character-

ize hard geometric objects in scenes which is an extremely useful property in removing noise from objects. Also low α indicates occlusions where there are multiple objects in a scene . Thus along modulus lines with low α we can separate two objects for identifying separate shapes. Because of the denoising capability we can also tune our algorithm to preserve only features of interest.

In terms of our multifractal parameters we can define α with

$$\alpha = -\frac{\log p}{\log l} \tag{4.2.2}$$

4.2.3 Fractal Dimension

The Lipschitz alpha has been related directly to what is known as fractal dimension[50]. Pentland [54]and others have shown that fractal dimension has shown promise in characterizing the texture natural objects from man-made and others have used fractal dimension to characterize the standard Broadatz texture classes. Unfortunately such texture characterization is arbitrary since it did not have a precise mathematical method but is instead discussed in terms of box counting dimension. Thus we must find another means of characterizing fractal dimension.

As it happens, the Lipschitz alpha is a difficult quantity to calculate experimentally for characterizing texture since it does not characterize individual points in an image but sets of points instead. To characterize individual singularity points in an image we turn instead to the Holder exponent. The Holder exponent h(x) of a distribution f at the point x_0 is defined as the greatest h so that f is Lipschitz h at x_0, i.e., there exists a constant C and polynomial $P_n(x)$ of order n so that for all x in a neighborhood of x_0, we have

$$|f(x) - P_n(x - x_0)| \le C|x - x_0|^h \tag{4.2.3}$$

4.2.4 Holder Exponent Revealed by the Wavelet Transform

We can use the wavelet transform to reveal the Holder exponent in a signal. Different wavelet basis sets have different ranges of permissible Holder exponents which they can reveal. To study this phenomenon as described by Arneodo[1], we slightly redefine the wavelet transform where

$$\psi_{b,a}(x) = a^{-\frac{1}{2}}\psi\left(\frac{(x-b)}{a}\right) \tag{4.2.4}$$

where a is a scale parameter $a \in \Re^+$ and b is a real valued space or time parameter. We now define

$$T_\psi[s](b,a) = \langle \psi_{b,a} | s \rangle_{L^2(\Re,\, dx)} = a^{-\frac{1}{2}} \int \psi\left(\frac{x-b}{a}\right) f(x)dx \tag{4.2.5}$$

where $\langle . | . \rangle_{L^2(\Re,\, dx)}$ is the scalar product in $L^2(\Re, dx)$. We can then derive that the local singular behavior $C|x - x_0|^{h(x_0)}$ of *f* around the point $x = x_0$ when the scale a goes to 0 for a given Holder exponent h(x). On the other hand , if *f* were C^∞ at x_0 on could prove that we would get a power law scaling exponent $n_\psi > n$.

$$T_{\psi}[f](x_0, a) \sim a^{n_{\psi}} \tag{4.2.6}$$

In other words, the maximum Holder exponent that one can achieve in a singularity spectra is equal to the number of vanishing moments of the analyzing wavelet. Thus for the first derivative of the gaussian as well as the Harr wavelet, both of which have one vanishing moment, we can achieve Holder exponents which are between +/-1.

4.3 2-D Singularity Characterization for Texture

Figure 4-6 shows the [39, 54] process of computing the Holder exponent on two dimensional images. This exponent is first computed by calculating the gaussian derivative decomposition of the image and then taking corresponding pixels across high pass images between successive scales. For each pixel we compute a graph of the slope of the wavelet coefficient magnitudes across scales. When the log (base 2) of the slope across scales is computed, each resulting slope is entered into a new image at the pixel position it represents.

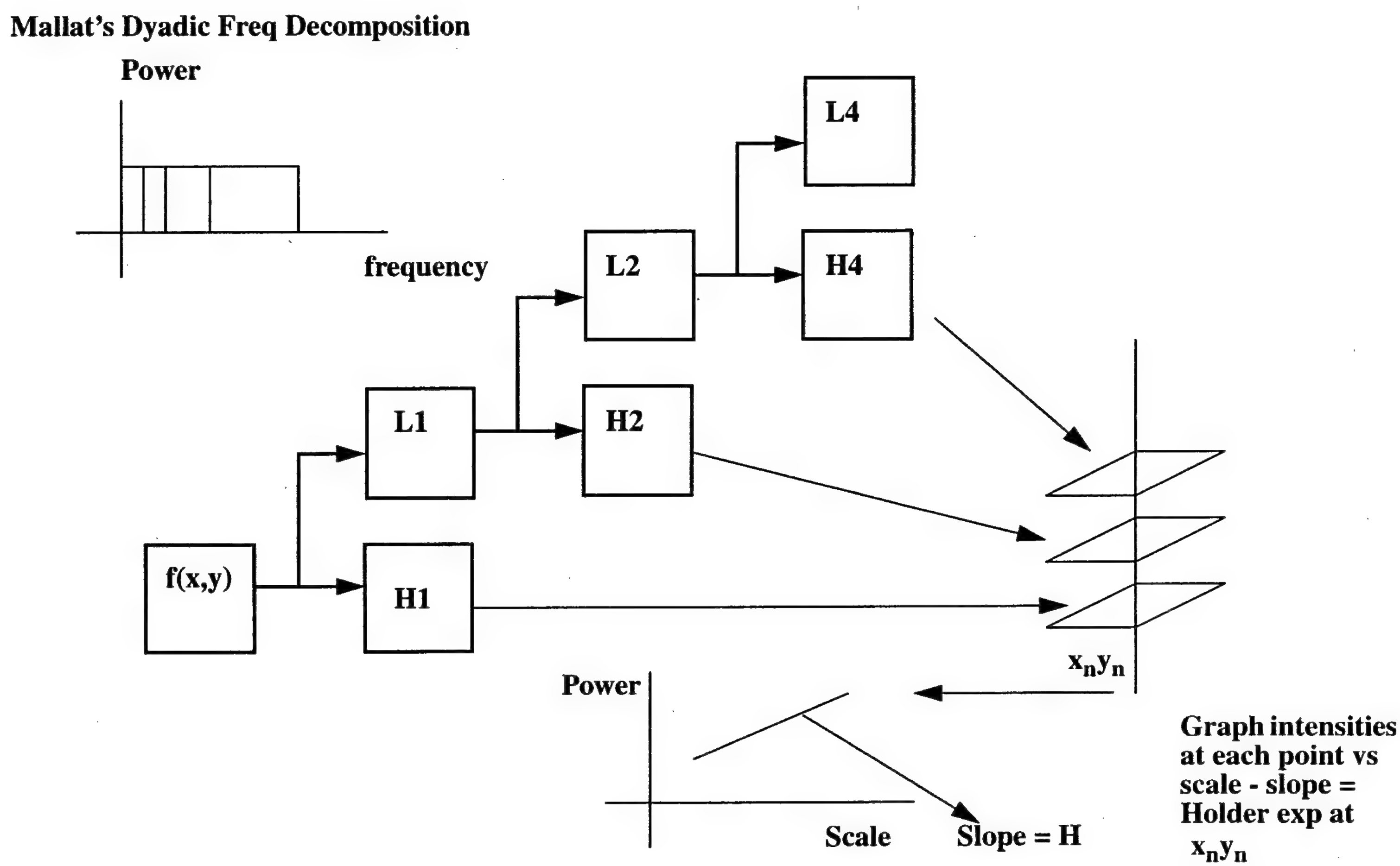


Figure 4-6 Singularity Spectra

The result is called the singularity spectra. An example of the singularity spectra is computed for the star image shown in Figure 4-7.

The result is that specific region extraction and object detection can be done with the singularity spectra by specifying Holder exponents in certain ranges. For instance, Holder exponents with negative values correspond to hard edges such as the edges of buildings in an image whereas positive Holder exponents correspond to naturally occurring phenomenon such as vegetatation, or cloud formation.

In the image of Figure 4-7 we are able to extract the shape of a planet from a very noisy saturated background simply by the fact that the planet has very hard fixed edges in the range from -.21 to -.2. If we were to use a different basis set having two vanishing moments such as the mexican hat function, we could reveal in a broader range between -2 to 2. However, this basis set also reveals many other unwanted features in the image for a similar alpha range so its detection capabilities are severely limited. The singularity spectra are an extremely useful tool in many applications [39, 54] including synthetic aperature radar analysis, downlooking satellite surveillance, and texture analysis. Recently similar analysis has been performed to characterize Broadatz texture classes as well.

Original Planet Image in Noisy Background

Planet Detected in Alpha Range -.21 to -.2

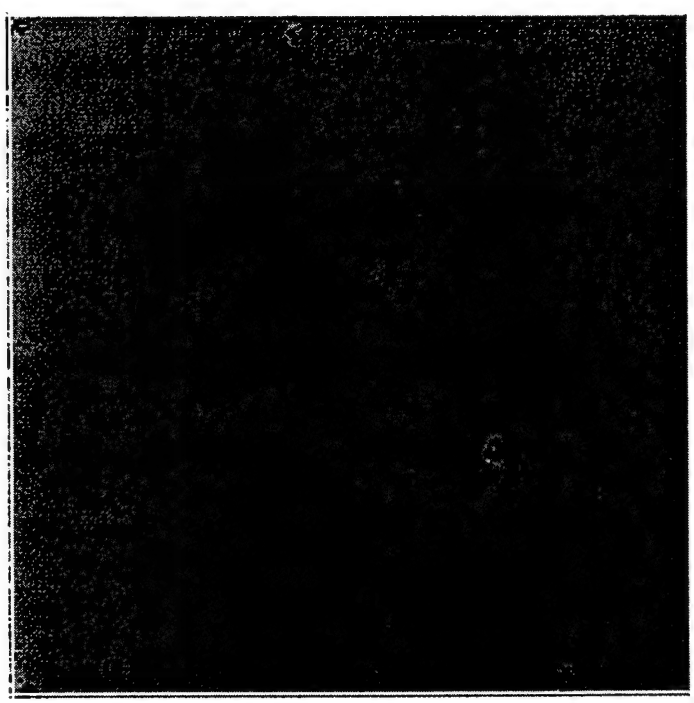


Figure 4-7 Planet Detected in Singularity Spectra

4.3 Analogy of Multifractal Global IFS to Local IFS

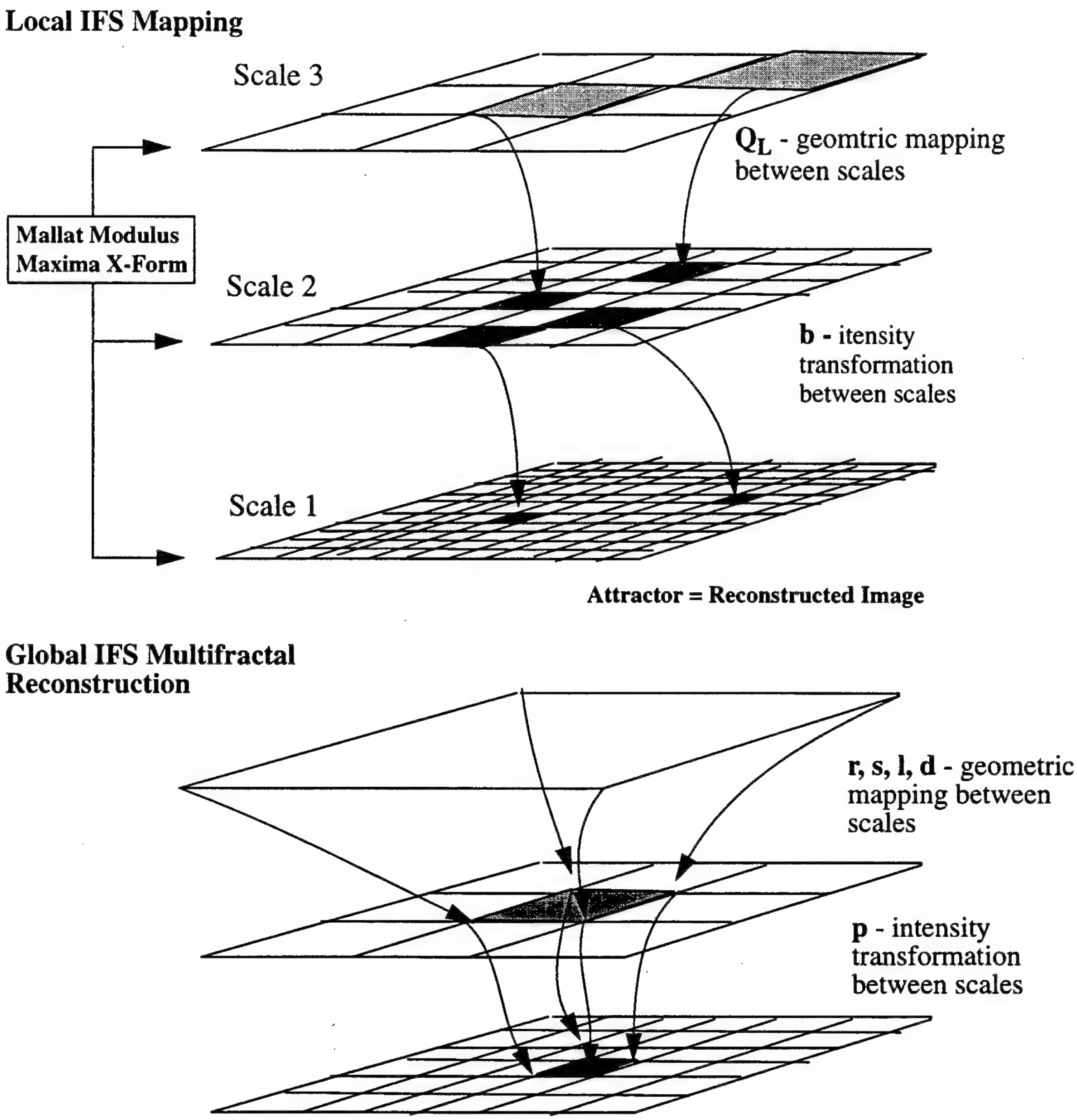
Now with the concept shape and texture as they apply to the multifractal we may now better understand some of the concepts of the local ifs which has been used to compress imagery. There are essentially three parameters needed for local fractal reconstruction as was discussed in Jaquin and Fischers's techniques and indicated in Equation 2.3.1 indicated below as Equation 4.1.1.

$$f(x) = Tf(x) = U_L f(x) + b = Q_L f(2x) + b \tag{4.3.1}$$

These parameters are spatial location and orientation symbolized by the spatial transformation Q_L analogous to the displacement parameters r and s in the multifractal case of equations 4.1.14 and 4.1.15. Note, however that the local IFS Q_L represents a block based operation between two scales where the r and s parameters represent a point operation between all scales. Secondly there is b the intensity transformation from range to domain block, analogous to the probabilty p in the 1 dimensional multifractal case. Again realized that b is an intensity transformation between two scales where p is an intensity transformation associated with r and s between all scales. Finally, the scale separation parameters, l and m are a power of 2 since we are using Mallat's dyadic decomposition.

Note that the local IFS technique has more flexibility than the global case since the mapping between frequency scales can be arbitrary where the global case has a fixed relationship between all scales. Thus the global case applies to only certain naturally occurring fractal objects such as the IFS fern or Sierpinsky Triangle. Note, as previously stated, it is not feasible to draw a direct mathematical relationship between the multifractal and local IFS parameters since the multifractal parameters represent pointwise transformations and the global ifs parameters represent block transformations. Nonetheless, the analogy will be useful when studying shape and texture

relationships later in this Chapter 5.



4-8 Local vs. Global IFS Mapping

Thus now in retrospect we can see the implicit relationship between shape and texture in the multi-fractal model. Recalling that the multifractal maps the lower frequency information to reproduce the higher frequency information in an image, lets assume we have an object with a

given shape outline in an image. This object may or may not be generated by a fractal process. Whatever the case, the object may also have some form of texture associated with it. If we assume an implicit multifractal equations associated with the object, then the gross edge outlines of the object are iteratively mapped back onto the object to reproduce its internal detail and texture. If so, the wavelet will will have a unique mapping from shape to texture. If the object is not fractal then the mapping will be an approximation. Thus, we are now prepared to develop a fractal wavelet method which has the basic elements of shape and texture as its fundamental components.

Chapter 5

Fractal Wavelet Method

Contents

5.1 Encoder

The basic principle of the fractal, as we have seen in preceding sections, is that of an object whose shape resembles that of its smallest component and there is a spatial mapping which defines the relationship between the objects overall shape to the smallest component shape. Now with Mallat's modulus maxima technique we have a means of defining shape at any given scale. The multifractal mapping represents the transition from an objects shape to its texture. This texture can be defined in terms of fractal dimension which in wavelet theory is the decay in energy of wavelet coefficients across scales. With this conceptual model we define the wavelet fractal method of image compression.

The first thing to remember when designing a fractal compression algorithm is that few objects that occur in imagery are generated by a natural set of fractal equations. Thus we are

forced to approximate objects in imagery with the fractal model. For this reason we will choose the local iterated function system to reproduce a given part of an image rather than a global iterated function system since the local iterated functions system has more flexibility in its construction.

Looking at all models of fractal encoding we realize that the trend is to reconstruct an object from its low frequency components to its high frequency components. In the wavelet fractal case this method approximates wavelet coefficients across scales. This model fits with the fractal method since low frequency corresponds to large scale. We also know that the Mallat multiresolution decomposition happens in diadic scales which, in the Haar basis set case, corresponds to blocks which are dimensions are powers of 2 in size. Thus we shall keep with this framework for the Gaussian derivative basis set. Putting this in the context of the wavelet transform we re-write equation 2.3.1 by inserting equation 3.4.24 as:

$$W_{2^{j-1}} f(x,y) = Q_L W_{2^j} f(x,y) + b \tag{5.1.1}$$

This equation is analogous to equation 4.1.9 for the global ifs multifractal case. Thus we build our reconstructed image from the low frequency or large scale images first and then eventually reconstruct the final image. Note that his process use the wavelet decomposition to explicitly separate scales by frequency. Thus for a given block size we only have information that is fits that particular scale.

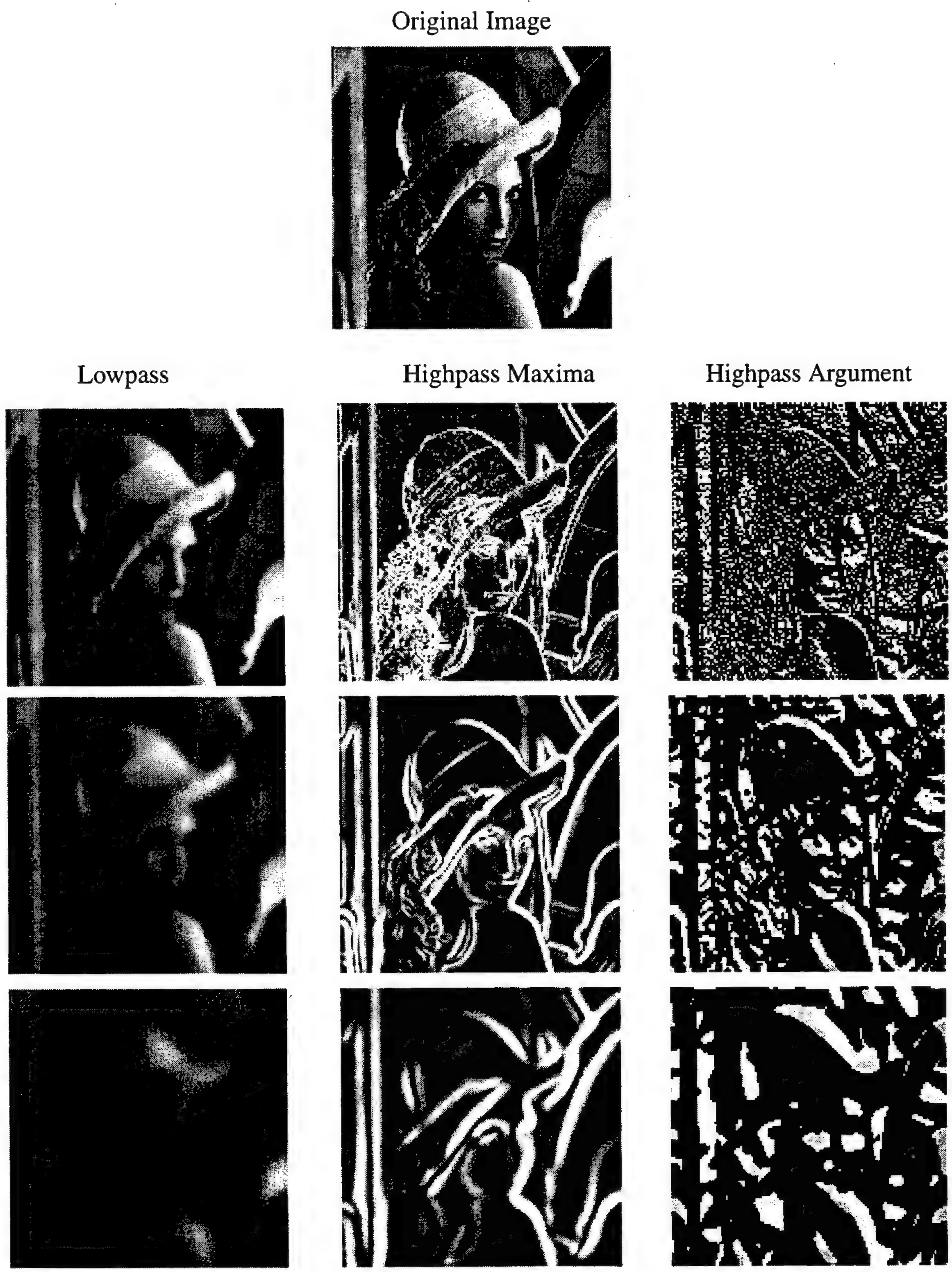
5.1.1 Modulus Maxima Decomposition

The first step in the wavelet fractal method is to reduce a given image by the Gaussian Derivative modulus maxima technique and obtain the wavelet scales W shown in equation 5.1.1. We recall that the gaussian derivative modulus maxima technique in two dimensions decomposes an image into three parts. A lowpass image defined be equation 3.4.19c as

$S^{d}_{2^{j+1}} f(x, y)$ the modulus maxima highpass magnitude $M_s f(x,y)$ and argument

images $A_s f(x,y)$. The results of this decomposition are shown for three scales on Lena below. It is important to note that in the Mallat decomposition that the scales are not subsampled as in standard wavelet decomposition so for range to domain block matching we subsample the higher scale by 2 as in standard fractal compression with no averaging process since each scale is naturally filtered by the wavelet.

Another by product of the Gaussian derivative wavelet decomposition is that it naturally organizes the features in an image for matching thereby restricting the domain to range blocks that get chosen for a particular scale, and thus makes the domain to range block matching process both accurate and fast. This is a significant improvement of Jaquin and Fischer's techniques of fractal compression.



5-1Modulus Maxima Decomposition Scales 1-3

5.1.2 Domain to Range Block Matching

To find a mapping Q_L from shape to texture for a particular object we need an efficient mechanism for determining a mapping. If we recall from section 2.3 range and domain blocks in traditional fractal encoding are obtained by subsampling the image and then matching each range block to every possible domain block in an image. Needless to say this encoding process takes an extremely long time and is one of the major drawbacks of traditional fractal encoding.

5.1.2.1 Range to Domain Localization Based on Shape

If we recall the concept of a fractal we remember that it is the same shape no matter what scale the user observes. Looking at the 2-dimensional Gaussian derivative basis set we see that the modulus maxima lines give us a natural indication of the object's shape as was indicated in Chapter 4. Thus, if we have a natural fractal object, we should be able to map the shape of the object into its smaller details via our local ifs block transformation process. This mapping process also typically follows the natural cone of influence of the wavelet scale decomposition of objects. [20] We will see how this process also speeds up encoding since it reduces the domain block pool and thus the matching time.

To begin the object based compression process, we must determine the boundaries and interior of an object. The modulus maxima values provide the natural boundaries of objects in a given scene for a particular scale. After choosing a scale at which the object of interest is located we take the modulus maxima values associated with that object and chain code them together based on their magnitude and gradient direction.[44] Once we have a set of modulus maxima values associated with a particular object we determine if these values form a closed curve. If they do we fill the interior and this becomes an object mask. If not we connect the ends of the curve with a straight line and then fill the interior of this curve for the mask. The process is shown

below in Figure 5-1.

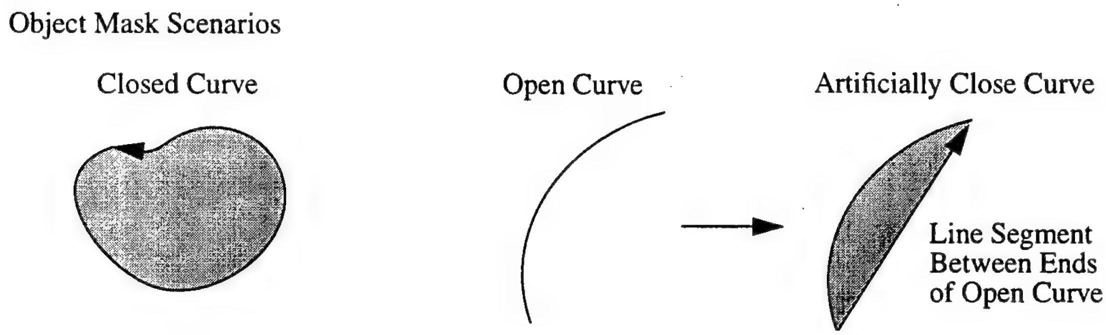


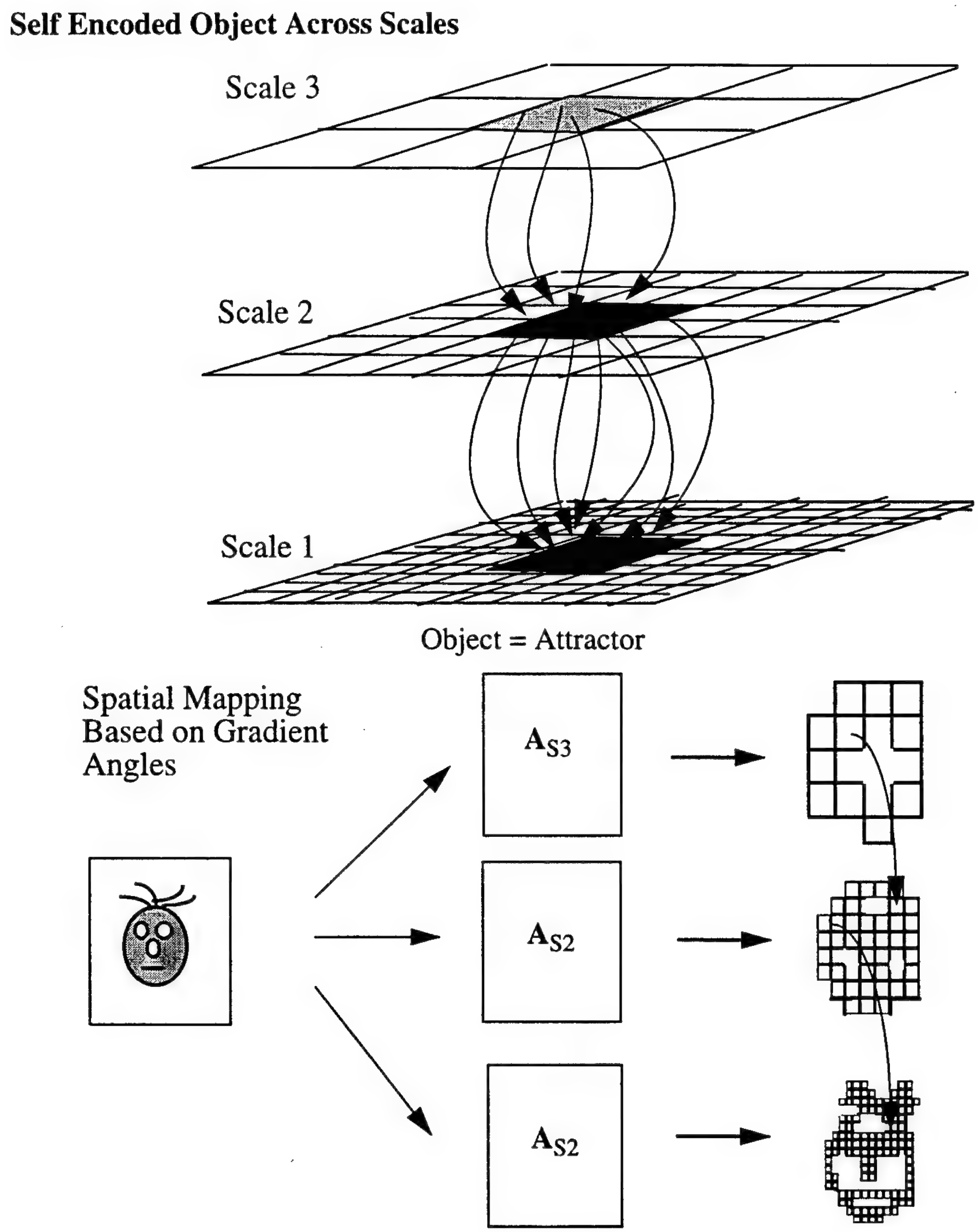
Figure 5-2 Object Shape Mask

If we encode using the object shape mask we select the range and domain blocks associated with that particular mask. This process makes intuitive sense, particularly if the object is generated by a self similar process such as a fern or a cloud. The modulus maxima mapping process in the case of a natural fractal would simply define how the object iteratively maps onto itself. Unfortunately since we are dealing with non-fractal objects in most case, and our block mapping procedure is restricted to dyadic scales with linear, square block transformations we can only approximate most natural fractals. Thus the block mapping process is at best a compromise between ease of implementation and exact representation.

An additional advantage of the self encoded object is that we can decode it individually apart from the overall image. This fact will later be extremely useful for image and video editing purposes. Also it offers the new possibility of object scalablity to reduce the overall bit rate of data transmission for a compressed file.

In some cases when objects are inherently non self similar restricting the domain block pool can result in reduced reproduction quality of an object. In the case where the best possible accuracy is required we can still fall back on selecting range and domain blocks from any

part of the image but in most cases this is not necessary. The object domain pool often gives results as good as those for the domain pool of the entire image. It also turns out that we can still save on bits in our encoded representation of the image with the b offset parameter since in the case of on object with constant texture the mapping between scales is the same and within an object we select blocks which have approximately the same b parameter. Also by having range and domain blocks naturally organized by modulus maxima lines we have a natural means organizing compressed information by shape and as we shall see texture. This fact is a significant improvement over the zero-tree method since it combines shape information as a natural part of the compression process. Examples of self encoded objects are featured in Appendix A.



5-3 Range to Domain Self Mapping

Thus, a direct analogy between the multifractal analysis and wavelet fractal encoding method may be drawn. The the object instead of the entire image is the attractor. The mapping between scales to reproduce a given object is simply designed to minimize distortion in the

reproduced object and not necessarily reveal the underlying physical process. However, knowledge of this technique improves encoder performance by speeding up the block matching process and reducing the bandwidth of compressed information if a user wants only a specific object from a scene rather than all the objects and background.

5.1.2.2 Modulus Maxima Gradient Matching

To simplify the process of range to domain block matching thus finding the mapping for Q_L in equation 5.1.1 we classify the range and domain blocks by summing the blocks modulus maxima magnitude and angle parameters since these gradient magnitude and angles are accurate indications of energy and direction within each block. Jaquin used a similar procedure in his classification of blocks in traditional fractal encoding. [37] by applying the centroid operator to each block. Now with Mallat's wavelet decomposition energy direction is already indicated as a natural part of the process. This block summation procedure is described for both angle and modulus values as is shown in equations 5.2.1 and 5.2.2. Note that Norm is the number of nonzero modulus or angle values in a block

$$\tilde{m}_{2^j} = \frac{\sum_{k=1}^{n} \sum_{l=1}^{n} M_{2^j} f(k,l)}{Norm} \tag{5.1.2}$$

$$\hat{a}_{2^j} = \frac{\sum_{k=1}^{n} \sum_{l=1}^{n} A_{2^j} f(k,l)}{Norm} \tag{5.1.3}$$

We then apply this same block matching procedure to the possible domain blocks at each scale. For domain/range block matching processes we use the difference of both angle and modulus values between successive scales to achieve domain to range block matching. The block pair with lowest combined *mdif* and *adif* are selected as the best matched blocks rather than recursively differencing every pixel in every possible range and domain block.

$$mdif = \left|\hat{m}_{2^{j-1}} - \hat{m}_{2^j}\right| \tag{5.1.4}$$

$$adif = \left|\hat{a}_{2^{j-1}} - \hat{a}_{2^j}\right| \tag{5.1.5}$$

Thus the operation of block classification thus becomes a lookup table procedure rather than an exhaustive matching process. Both range and domain block position in the image are stored . The block rotation value is also determined by applying the appropriate flip that makes the block gradient angles match most closely.

Until recently one of the greatest drawbacks to fractal and fractal wavelet compression has been overwhelming number of computations necessary to compute the range to domain block mappings since the process was performed by least mean square differencing for every range to domain block pair for multiple scales. Our matching technique dramatically reduces the number of calculations to compute a compressed image. The number of calculations is $O(N^2\log(N))$ (a given image has N^2 pixels) which is a significant improvement over existing fractal wavelet techniques[9] which can be as high as $O(N^3\log(N))$ in complexity. Thus this method is a significant improvement over the Davis and Rinaldo/Cavagano[9] methods of fractal wavelet compression.

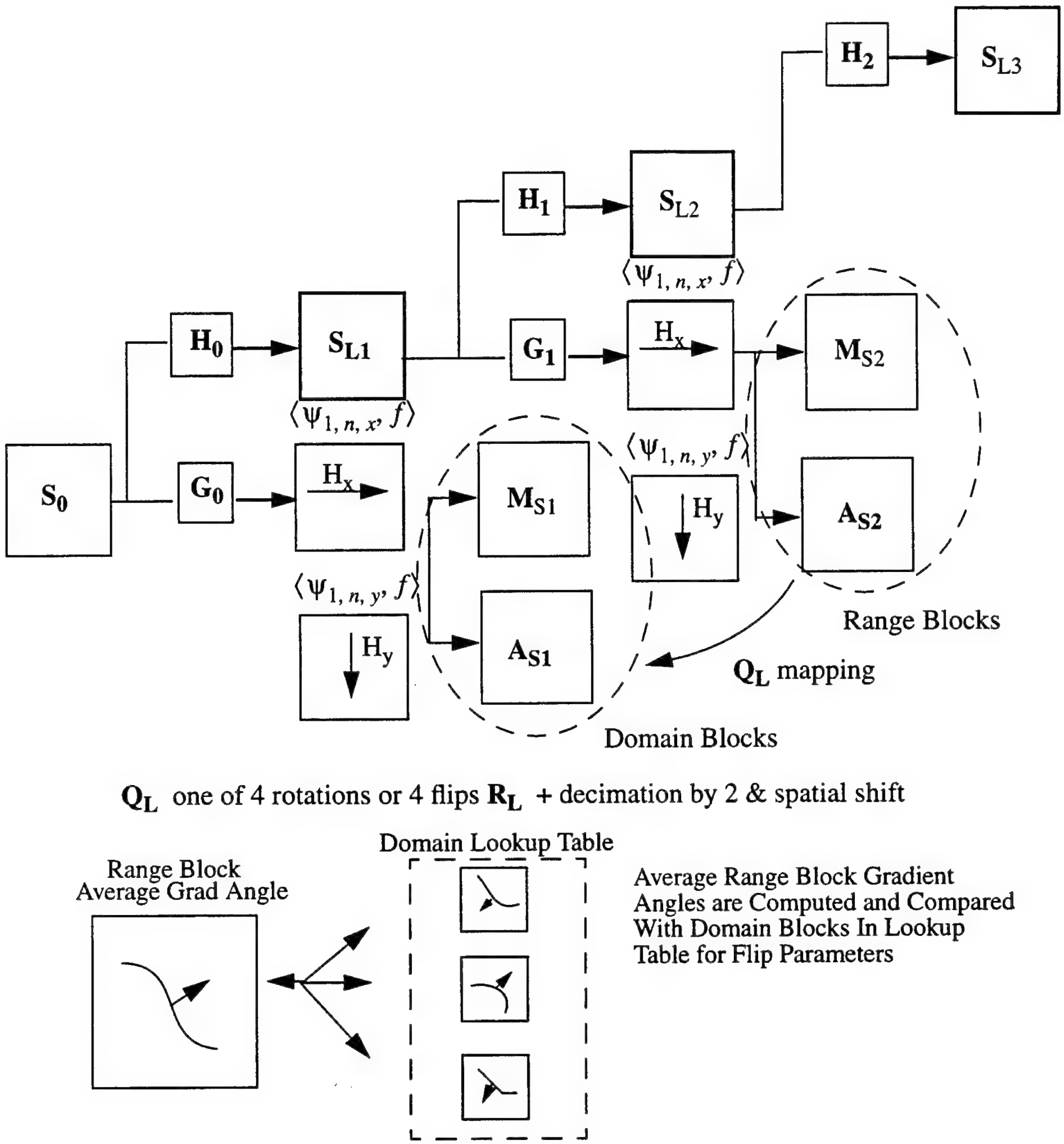


Figure 5-4 Gradient Angle Matching

5.1.3 Intensity Offset Determination

In order to determine the intensity offset or b parameter from Equation 5.1.1 between two scales we find the range and domain block pairs positions determined by gradient angle matching

and then subtract the lowpass blocks which correspond to these positions. Thus within a given range block where the dimensions of the domain and range block are range {x:x1<x<x2} and {y:y1<y<y2}.

$$b = \sum_{x=x1}^{x2} \sum_{y=y1}^{y2} \left(S^{d}_{2^{j+1}} f(x,y) - S^{d}_{2^{j}} f(x,y) \right) \tag{5.1.6}$$

This strategy makes sense since at each stage of the reconstruction process we are using a given lowpass image to reconstruct the information in the next highpass band. The spatial position information via range to domain block matching is computed with the highpass modulus maxima bands and all parameters necessary to complete the calculation of equation 5.1.1 are present.

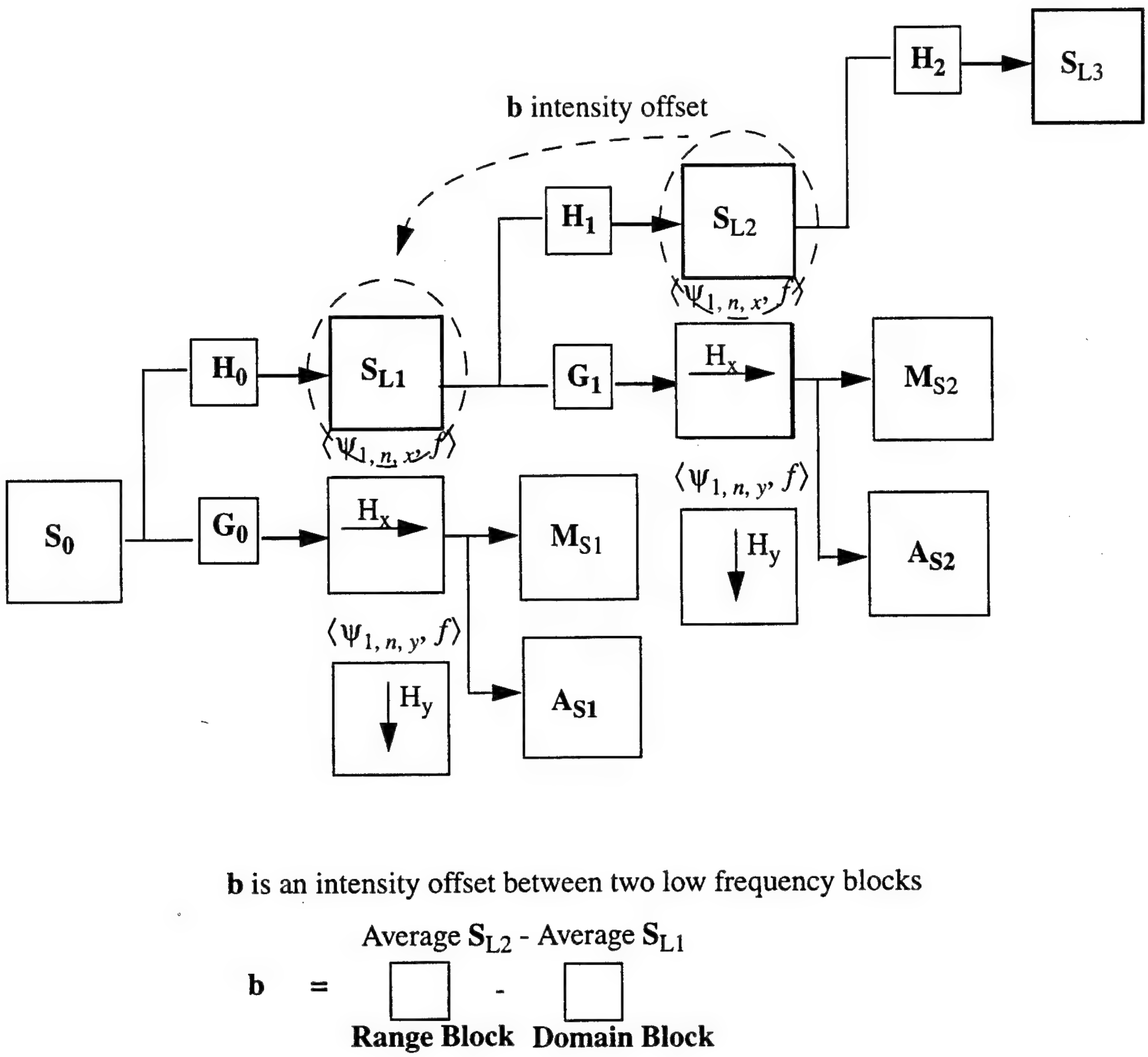


Figure 5-5 Dyadic Maxima Fractal Block Matching

5.1.4 Zero-Tree pruning

Essentially the fractal mapping process amounts to approximating successive wavelet frequency scales with information from previous scales. This decay across scales is known as texture in the sense of fractal dimension. Thus as was stated before the fractal mapping constructs its texture from its shape. Shapiro noticed that the natural decay of fre-

quency between scales can be exploited to optimize the compression process by eliminating wavelet coefficients that fall below some threshold.

This pruning of wavelet coefficients can be adapted for the fractal wavelet compression method. We can show that using the modulus maxima to determine the range to domain mapping can result in a natural quadtree pruning process. Since the edge detection process naturally compresses the spatial information to the maxima locations and these maxima are in turn thresholded, blocks with maxima energy falling below a certain threshold are naturally eliminated thus improving the overall compression process. Others such as Davis [20, 59] have included a rate distortion optimization routine for this process although our technique is computationally faster and has comparable compression performance with rate distortion optimized code. Once Zero-Tree pruning is performed we run-length encode range and domain information at each scale.

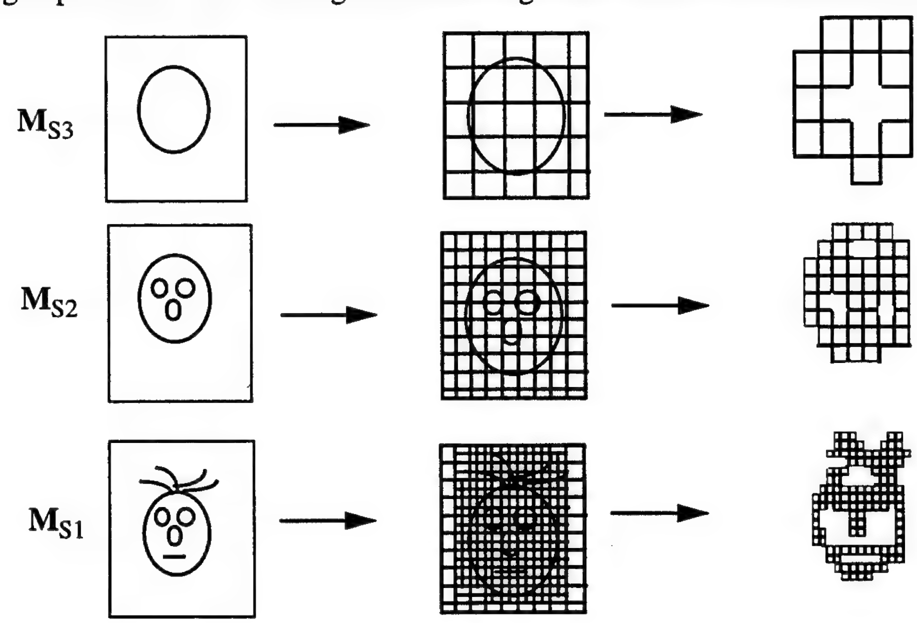


Figure 5-6 Dyadic Maxima Zero Tree Pruning

5.1.5 Color

For flexibility in compressed domain color storage and compressed domain object manipulation we transform RGB images into YRB images in the compressed domain corresponding to the follow color transformations. For luminance we have the transformation T_c

$$Y = 0.177R + 0.831G + 0.011B \tag{5.1.7}$$

To return from luminance to RGB we have T_c'

$$G = 0.135B + 1.23G + 0.22R \tag{5.1.8}$$

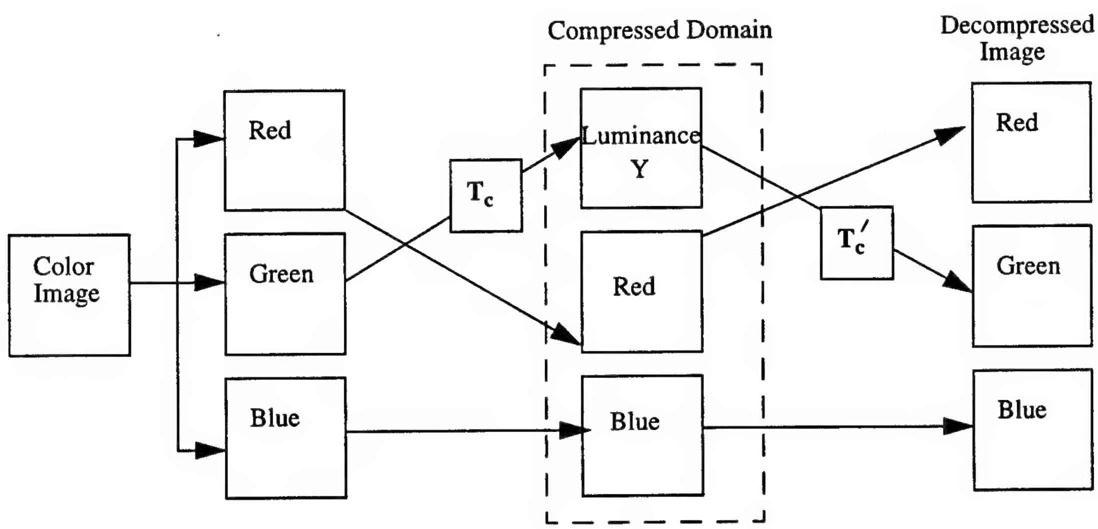


Figure 5-7 Color Image Transformation

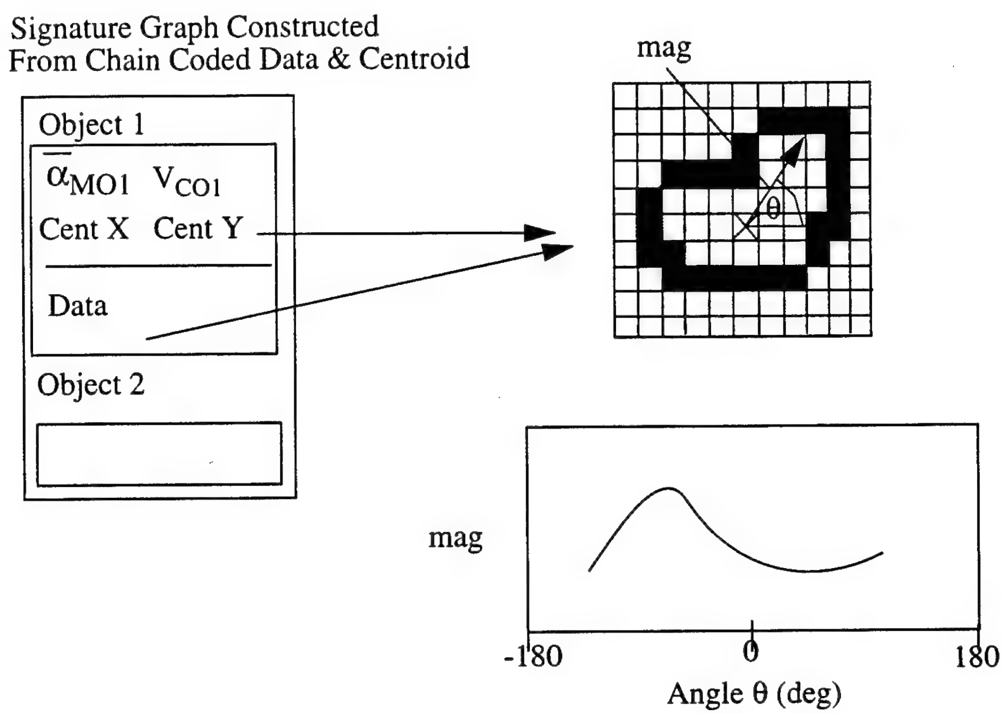
Overall we decompose the image into three separate color bands and each one of which has a separate multiresolution modulus maxima analysis performed on it. An overview of this entire process is featured in Appendix A.

5.2 Image File Organization

With the great amount of information obtained as the result of the wavelet analysis process we must find an efficient way to characterize this information in the compressed domain. We will now develop a means of organizing shape, texture, and color in the compressed file so it is easily accessible to the user in its compressed form.

5.2.1 Object Edge Blocks & Signatures

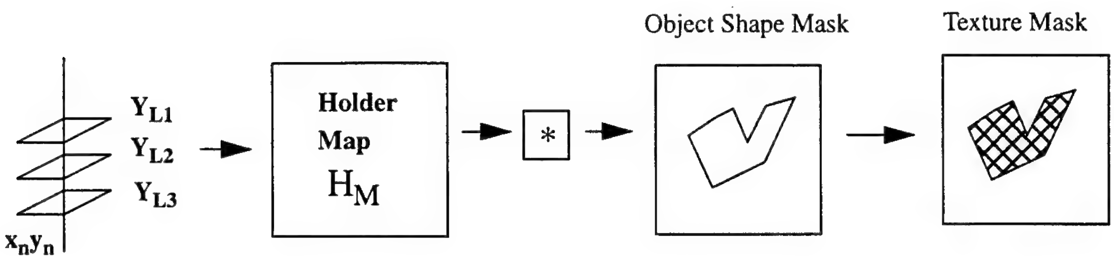
Because the range blocks are chosen basis on the modulus maxima edge formulation and their positions are stored we can isolate edge blocks from interior blocks in the compressed domain. From these edge blocks we can recover the shape outline of each object through a technique known as a signature. A signature is computed by finding the geometric center of an object and then making a plot of the angle vs. the distance to the edge from the geometric center. This technique is extremely useful since it is invariant to scale and can be adjusted for rotation in 2 dimensions. [32] The only requirement for this technique is that it have reliable edges information. Figures show the signature matching process. Signatures may be matched by simple LMS differencing algorithms or more sophisticated correlation algorithms for the case of rotated objects. The following figure shows the signature creation and matching process. Example signatures are shown in Appendix A.



5-8 Signature Creation and Matching

5.2.2 Object Texture Feature Vectors

Not only can the modulus maxima shape mask be used to restrict domain block search but it can also be used to segment relevant texture information about an object. To define the texture within an object we create the object texture mask multiplying a binarized copy of the shape mask with the Holder exponet map of an image. The interior of this mask corresponds to the profile of the texture within a given object.



5-9 Object Texture Mask

Within a texture mask we can compute the average texture as shown below. This is known as the texture or Holder vector for that object.

$$\bar{H} = \frac{\sum_{x=0}^{N} \sum_{y=0}^{N} H(x, y)}{N \cdot N} \tag{5.2.1}$$

5.2.3 Object Color Feature Vectors

Also the Red, Green and Blue partitions of a color image may also be segmented using the object shape mask as is shown below. Thus we can associate color with given objects in an image.

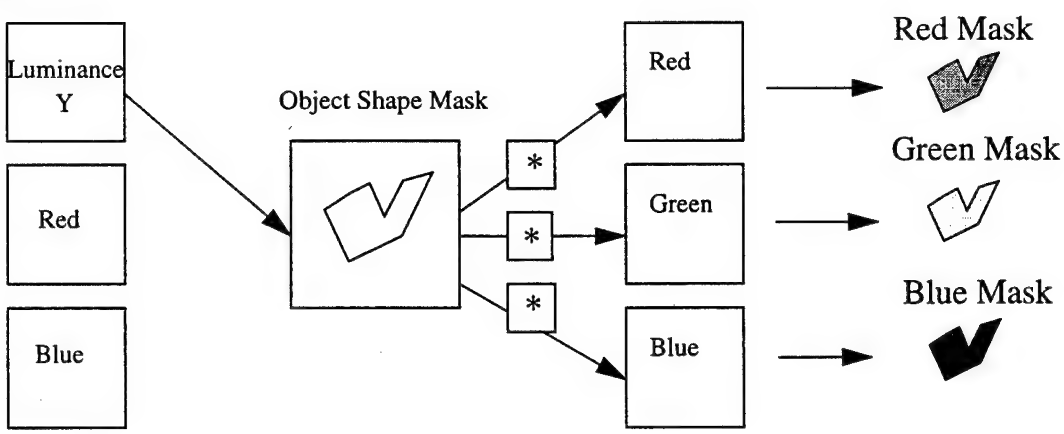


Figure 5-10 Color Mask

Color feature vectors V_C and the associated average colors $(\bar{R}, \bar{G}, \bar{B})$ within an objects mask are described below:

$$\bar{R} = \frac{\sum_{x=0}^{N} \sum_{y=0}^{N} R(x, y)}{N \cdot N} \tag{5.2.2}$$

$$\bar{G} = \frac{\sum_{x=0}^{N} \sum_{y=0}^{N} R(x, y)}{N \cdot N} \tag{5.2.3}$$

$$\bar{B} = \frac{\sum_{x=0}^{N} \sum_{y=0}^{N} R(x, y)}{N \cdot N} \tag{5.2.4}$$

$$V_C = (\bar{R}, \bar{G}, \bar{B}) \tag{5.2.5}$$

5.2.4 Object Compositing by Feature Vectors

Quite often edges that define objects do not form closed curves because of variations in image intensity, occlusions or other image artifacts. Thus signatures do not always reflect the true shape of an object. To combat this problem we can group objects together. Objects with similar feature vectors can be grouped together as one object if their relative positions are near each other. This operation can be user defined or automatic depending on the desired information in the composited objects. An example of this process is shown in Appendix A

Such object compositing can be useful for providing more accurate signature information since quite often there are gaps in the chain coded block edges which can be filled in with texture and color information as is shown below.

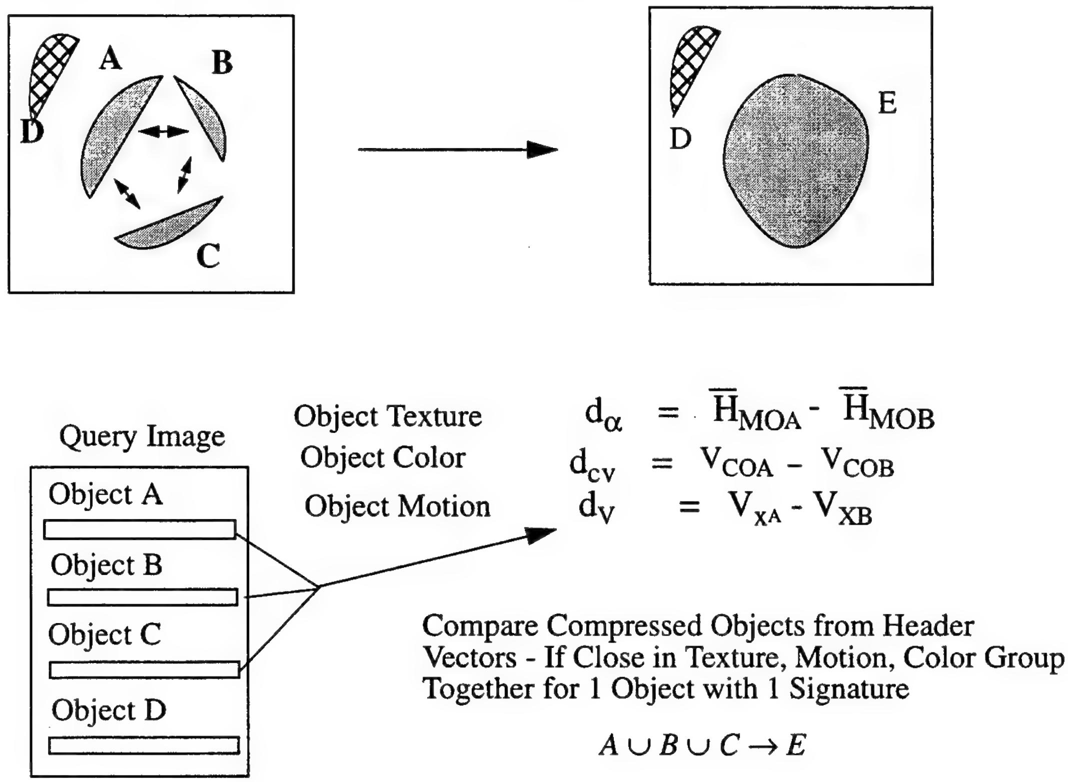


Figure 5-11 Object Grouping

5.2.5 Feature Vector Matching Process

To compare shape, texture and color of objects within the compressed file we need simply compare the associated vectors. This greatly increases the speed of compressed object database matching. The matching process is also quite straightforward since vectors can be matched by simple LMS differencing or correlation in the case of shape signatures. A graphical overview of the matching process is shown in the following

figure and Appendix A.

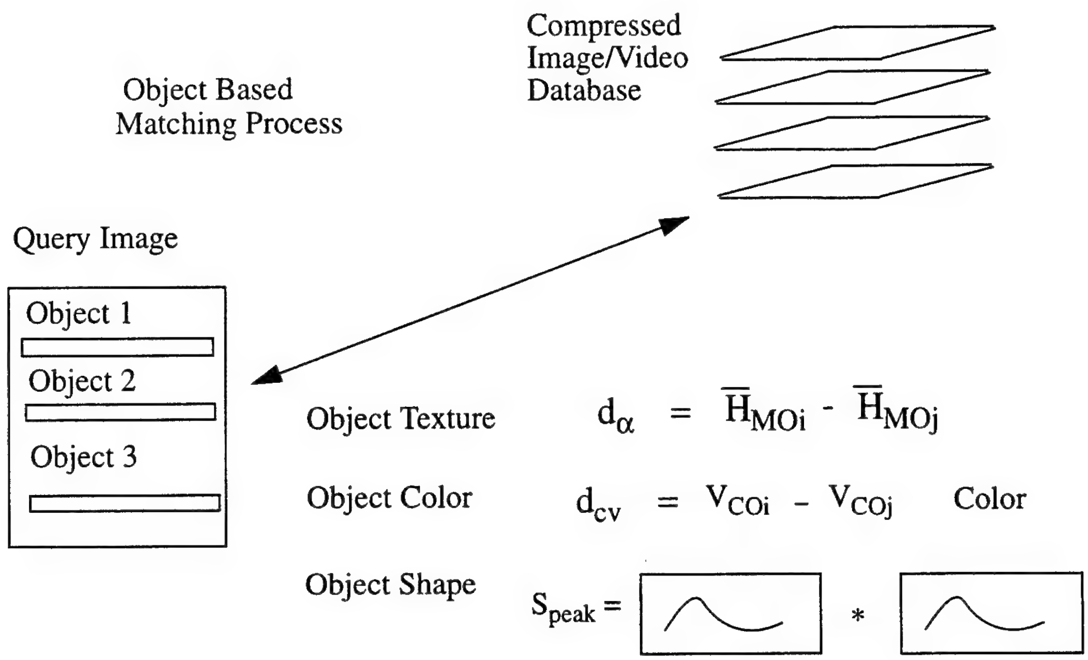


Figure 5-12 Signature and Vector Matching Process

5.2.6 Stored Image File Format

Combining the reorganized wavelet fractal compressed file information we have he object based wavelet fractal file. This file contains all the information necessary to interpret shape, texture, color, position and orientation of objects in a scene. Images can be rendered progressively, objects can be rendered individually, objects can be rendered in black and white or color, and every object has shape, color, texture, position, and orientation information associated with it. Thus such an image file is ideally suited for content based inquiry applications as well as interactive video editing and in an extremely low bit rate environment. Formulas for object moment calculations are shown in Appendix B. An example compressed image file is shown in Appendix A

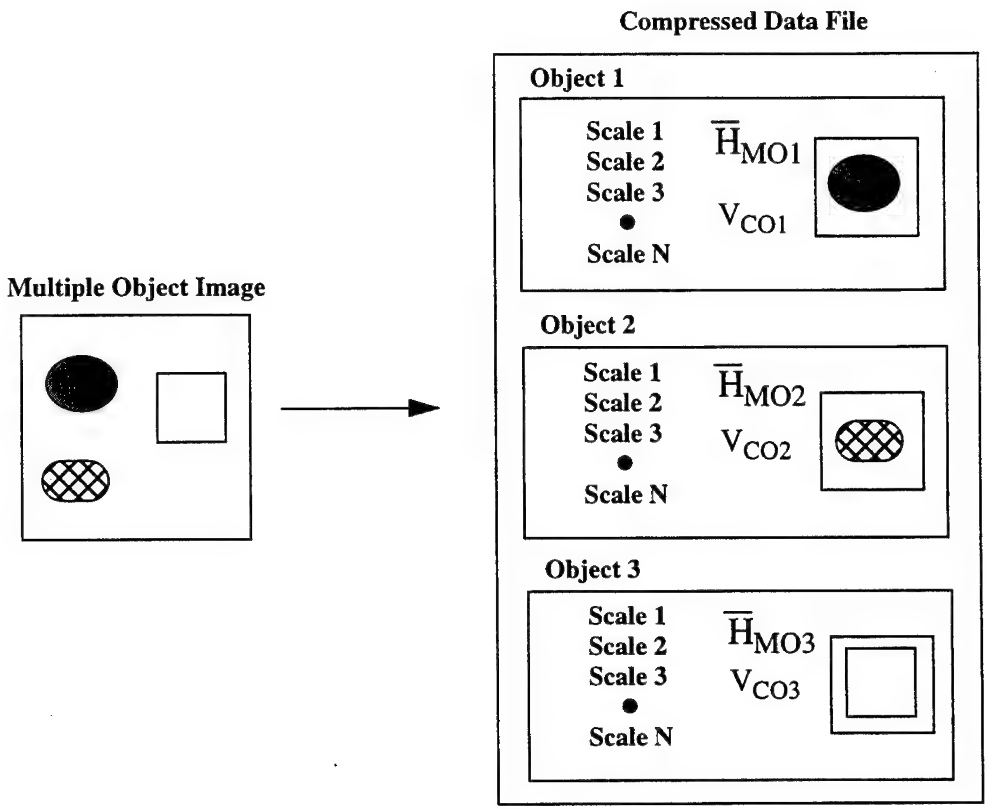


Figure 5-13 Image File Format

5.3 Decompression

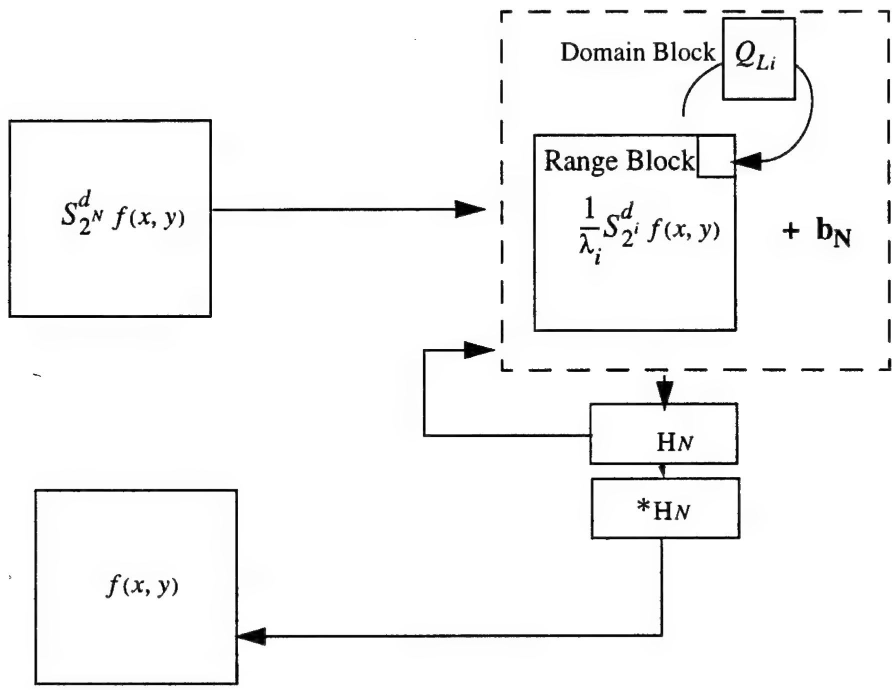
Because our frequency mapping proceeds from low to high frequency our reconstruction process simply proceeds by iterating our compressed parameters on the low pass image. For each scale the iterative procedure forms the approximation of the next higher lowpass image and is then lowpass filtered and the process is repeated. This process thus removes all blocky artifacts in the image while still revealing the image features. The above technique leads to a direct reconstruction method based on the fast decomposition process described in section 3.4.3

j = J
while (j > 0)

$$S^{d}_{2^{j+1}} f = Q_L \frac{1}{\lambda_j} S^{d}_{2^j} f * (\tilde{H}j{+}1, \tilde{H}j + 1) + b \qquad (5.3.1)$$

j = j - 1
endwhile

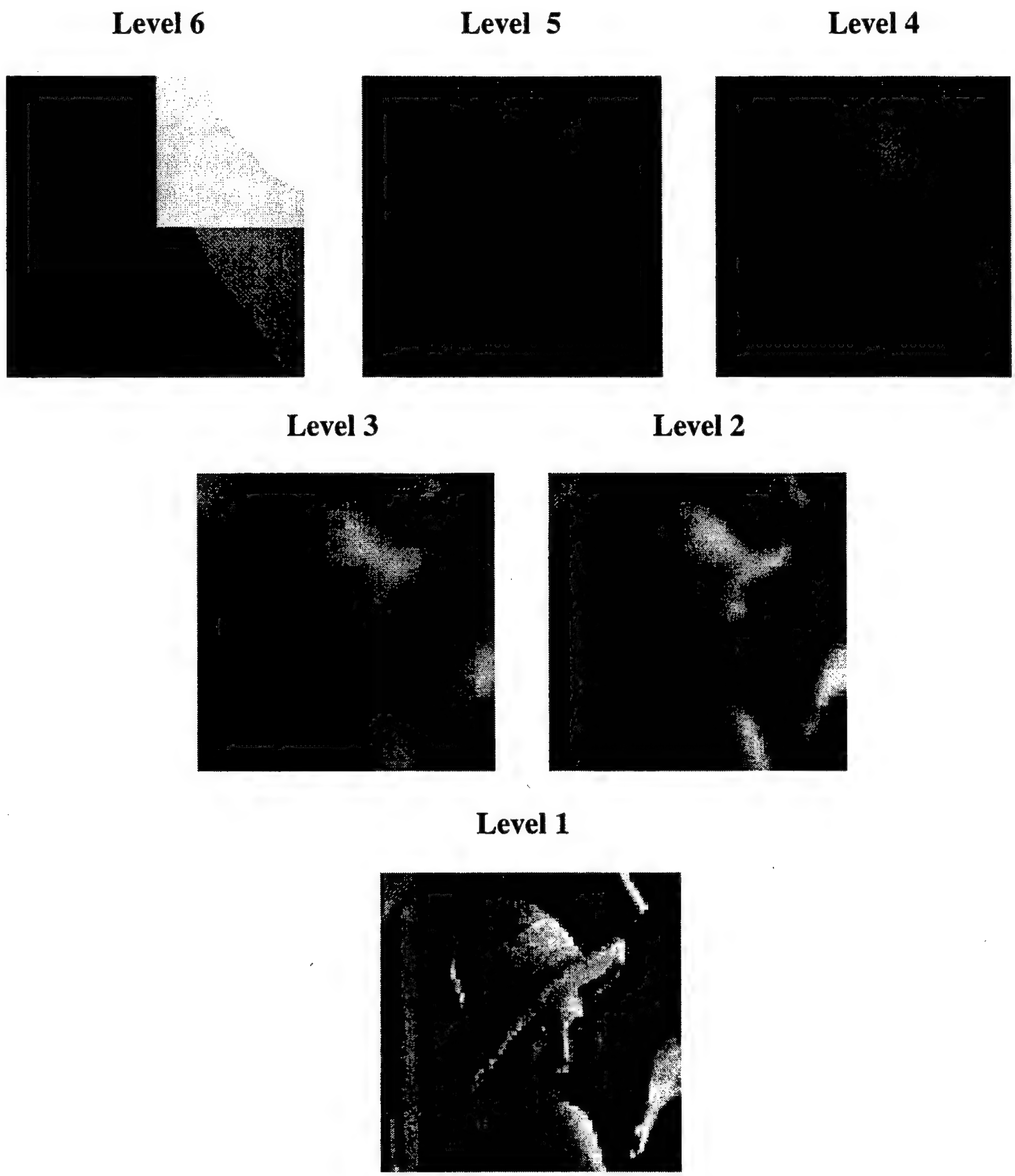
The stages of this reconstruction process are indicated graphically as follows.



5-14 Direct Reconstruction Technique

Thus in the compressed file our first step is to upsample compressed lowpass image and then use the stored fractal local iterated function parameters on it to reconstruct the original image building each new scale from the previous lowpass image. The image can be restored to any desired resolution simply by stopping the reconstruction process at a given scale. This iterative procedure is only order O(N) where N is the number of image blocks for all scales and thus can be performed real time with no special hardware. Thus this approach is a significant advance over the Mallat alternating projections reconstruction method because it is extremely efficient in its reconstruction

speed. The reconstruction stages for 6 scales are shown in the following figure.



5-15 Multiresolution Fractal Wavelet Reconstruction

Thus our compression process is progressive or spatially scalable and is well suited to network transmission. For low bandwidth systems lower resolution copies of the image are transmitted first so the user can see the image before the entire file is transmitted. This fact will be extremely useful in our video compression technique. We can see in the decompressed versions of Lena that in fact most block artifacts are removed from the image and we have a very high quality reproduced version of the original image. A graphical overview of the decompression process is featured in Appendix A.

5.3.1 Results

Figures 5-16 through 5-17 show the test result of transforming the images of Lena with the multiresolution transform in comparison with other popular compression methods. In figure 5-18 we see that the multiresolution technique has higher PSNR for a given compression ratio than pure wavelet. Our method falls slightly below zero-tree compression but in general follows the same compression numbers for Davis's spline wavelet fractal approach. At low compression ratios our technique is equal to or slightly below most of the compression techniques in PSNR. At 16:1 our wavelet fractal approach surpasses most conventional fractal and DCT methods and maintains a relatively flat PSNR curve where other methods tend to decrease in reproductive quality [9]. This is due to the fact that by filtering at each decompressed scale we remove artifacts which appear in linear transform techniques as well as fractal techniques when the number of coefficients or fractal blocks decrease.

Compress Ratio:	4:1	8:1	16:1	32:1	64:1
Bits/Pixel:	2.0	1.0	0.5	.25	.125
PSNR(db):	38.6	36.7	35.8	33.5	31.7

Figure 5-17 Lena at Various Compression Ratios Clockwise 1:1, 8:1, 16:1, 32:1

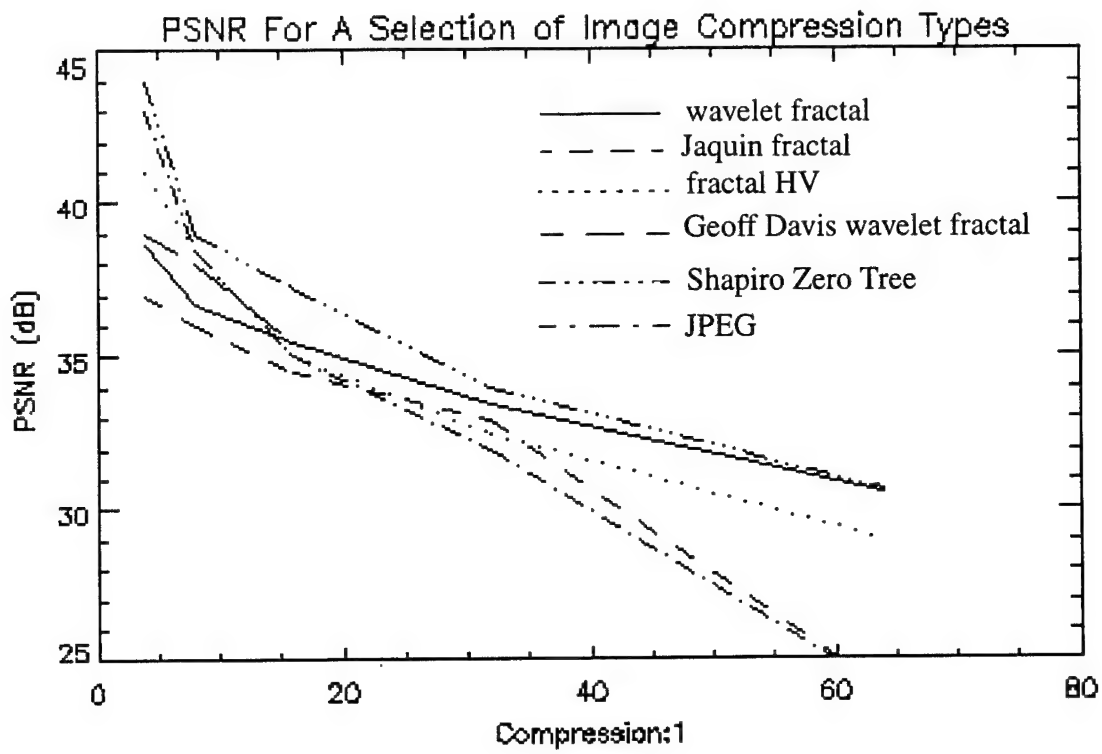


Figure 5-18 Comparison of Compression Ratio vs. PSNR for Several Compression Types

Chapter 6

Video Encoding

Contents

6.1 Object Based Video Encoding

As a result of the image compression method we have developed, we can apply the same principles to object oriented video coding. Most video motion estimation techniques rely simply on motion as a way to define objects in video. However, given our concept of shape- to- texture mapping we have a means of defining a video frame by its shape, texture, and color in one integrated approach and then using optical flow [61] to define object-based motion.

First [17] we know that for object-based coding to work effectively, and improve over strictly block-based coding, only a few objects in the scene are moving, object motion is dominant, and moderate, the moving objects cover 40-60% of the image area, and no camera motion occurs. For the general crossection of video we need to have a technique which can handle video favorable and unfavorable to object video.

6.2 Optical Flow

For video compression we have an effective means of detecting motion of objects in the video stream. Sundeswaran [67] has shown that one can compute optical flow from Mallat's edge detected imagery. If the image intensity function is represented by I(x,y,t), one can prove that the two components of the optical flow (Vx,Vy) satisfy the optical flow constraint equation 6.1.1.

$$\frac{\partial I}{\partial x}V_x + \frac{\partial I}{\partial y}V_y = \frac{\partial}{\partial t}I \qquad (6.1.1)$$

At a fixed time t, instead of solving the motion constrain equation for the image I(x,y,t), we can first smooth this image with the smoothing function θ(x,y) dilated by 2^j. This reduced the computational noise when estimating partial derivatives with finite differences. We thus have equation 6.1.2:

$$\frac{\partial}{\partial x}(I_t \otimes \theta_j)V_x + \frac{\partial I}{\partial y}(I_t \otimes \theta_j)V_y = -\frac{\partial}{\partial t}(I_t \otimes \theta_j) \tag{6.1.2}$$

This equation allows us to recover the normal component of the flow from the wavelet transform at the scale 2j Instead of computing this normal component at all points (x,y) we compute it only at the locations were the wavelet transform modulus is locally maximum. This saves significantly in computational complexity over traditional optical flow computation techniques.

Thus for each scale we compute the optical flow. Since each dyadic scale has a particular block size associated with it we take the average optical flow associated with each NxN block as we did the average argument and modulus maxima as is described by the following equations.

$$\overline{V_x} = \frac{\sum_{x=0}^{N}\sum_{y=0}^{N} V_x(x,y)}{Norm} \tag{6.1.3}$$

$$\overline{V_y} = \frac{\sum_{x=0}^{N}\sum_{y=0}^{N} V_x(x,y)}{Norm} \tag{6.1.4}$$

Optical flow associated with small blocks is associated usually with background texture motion such as moving water or camera panning where optical flow associated with larger blocks is associated with object motion

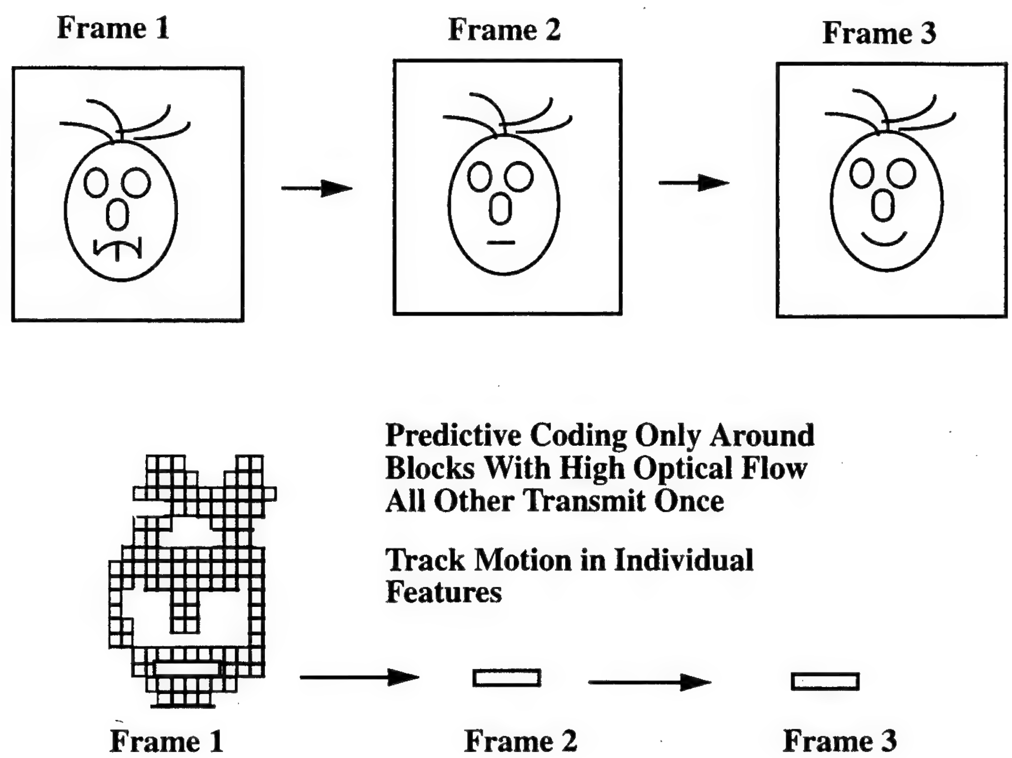


Figure 6-1 Video Sequence Coding

6.2.1 Frame Structure

In our current implementation we have only intra coded or I frames and forward predicted or P frames. Within this structure I frames are self encoded as normal image frames and optical flow is computed between an I frame and sucessive P frames. This model assumes there is no significant scene changes between two I frames.

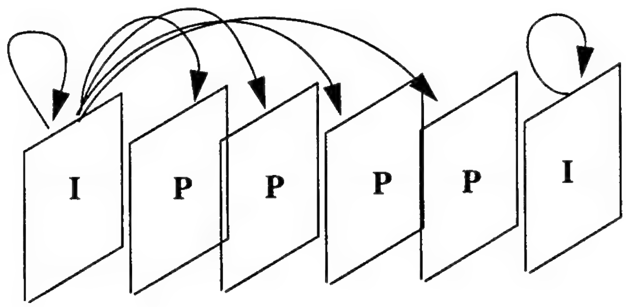


Figure 6-2 Frame Structure

6.2.2 Background

For static background characterized by average optical flow in blocks below a certain threshold we do not retransmit these blocks between frames. Moving background this is simply treated as an added object. Individual block motion vectors are transmitted to update background and all background blocks are recoded if motion is above a given threshold. Since this process is done in a multiresolution fashion, smaller detail changes may be added to update fine texture while lower resolution information remains constant. Uncovered background is totally recoded as if it were part of an intra frame, as are 3 dimensional changes in objects and new objects that enter a scene or scene changes. Such changes are characterized by optical flow above a given threshold which is set based on desired video quality vs. bit rate requirements.

6.2.3 Moving Objects

As presented from Chapter 4 we define shape at a particular scale j by taking the chain coded outline of an object at a particular scale and forming a mask. This wavelet defined shape technique is simply a by-product of the modulus maxima decomposition. We may apply this same technique to the optical flow information from the modulus maxima computation. A video object mask is created in the same manner as other objects. A binary shape mask is multiplied by the optical flow x and y images as is shown below. The average flow within this object thus character-

izes the average shape based motion of the object.

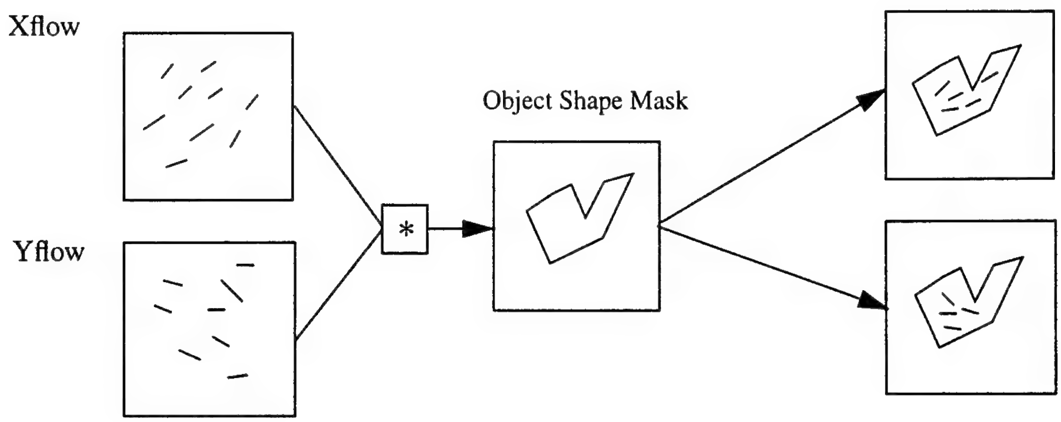


Fig 6-3 Motion Object Mask

This object based format can result in significant bit rate improvements over transmitting individual frames. The optical flow format, by concentrating motion into different frequency scales does an excellent job of characterizing motion which might occur within and object. For instance the external gears of a moving train would be characterized by one scale while the overall train engine itself would be characterized by another. Each would have its own motion but one object could be contained within the other at a different scale.

It also interesting to note that such block based optical flow analysis can give us rotation and three dimensional information about objects[5]. In a rotating object the edges will rotate faster than the center and thus depth may be derived via motion vectors. Also rotation may derived with concentric vectors around a central axis. A graphical overview of the video compression process is shown Appendix A.

6.2.4 Object Integrity

In order to determine whether an object has changed significantly between frames so it should be updated we can use the associate feature vector with the object. If the overall shape of the object changes we can compare 0th moment or signature profile. If the angle of the object has changed we can observe the 2nd moment or profile of the flow vectors at the edges. For 3-D rotation we can look at the flow at the edges of the object. If internal details of the object have changed we can look to average texture information or flow vectors at smaller scales. If there are significant changes the object is totally re-coded. Thus the object information for compressed domain query also improves the motion estimation process. Also compositing may be used to improve object quality. A pictorial example of this process is shown in Appendices D-2, D-3, and D-4.

6.2.5 Video Object File

The video object file has the same capability as the image file with the addition of an optical flow field to track individual object blocks for motion in P frames. In I frames all range and domain block information is transmitted. As in the image case three separate frames are encoded after a color transformation has been performed. One for luminance, one for the red frame and one for the blue frame. An example video file is shown in Appendix A.

6.3 Decoder

As with the fractal video system the decoder is efficient, needing no special hardware for near real-time decoding. Only those blocks which change and are not part of existing objects need to be retransmitted; thus the decoder only needs to update those blocks that completely change. Because the user can isolate self encoded objects these

can be transmitted individually so that video editing functions may be performed without special human intervention.

6.4 Scalability

This intra frame process to a new set of scalabilities in image transmission as shown in Appendix A . The first is spatial scalability - low frequency scales are transmitted first and then higher frequency information follows as new scales are added. Secondly we have temporal scalability with P frames being skipped of the video stream depending on the speed of the connection and the amount of temporal resolution needed. Also we have object scalability where only individual objects are transmitted or just their motion vectors. Finally, we have color scalability where black and white luminance information or full color can be transmitted depending on user requirements.

Appendix A show how the above system can be integrated into a distributed networked content-based query system. Thus we may manipulate individual objects by shape, color, texture, and motion in video sequences. We may database search on individual frames and post process video to modify content. The issue of how to design a network to optimize this search process is still an open topic but the evolution of such tools as JAVA for object oriented programming is greatly increasing the ease of this process.

6.5 Results

The results of our study are shown as follows. We attempted to maintain the same PSNR within the taxi sequence shown as the still encoded images of Lena. In this compression study motion was computed simply by computing the flow between an I and given P frame where I frames occurred every 5 frames. Future studies will examine in detail object oriented motion compensated video. Factors that decrease the compression ratio achieved include camera jitter,

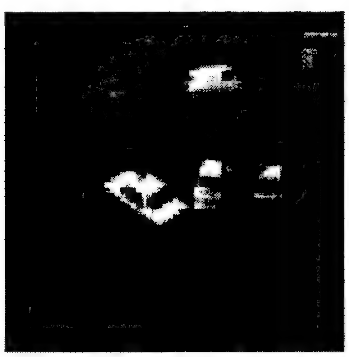
photon noise within pixels, and camera non-uniformity. Upcoming studies will focus on removing such image distortions to obtain even higher compression and lower overall picture noise. The prominent spike shown in these curves is the result of transition from an intra to inter coded frame as well as a large truck entering the scene and a car turning away from the camera in 3-D.

Taxi Sequence Frame 1

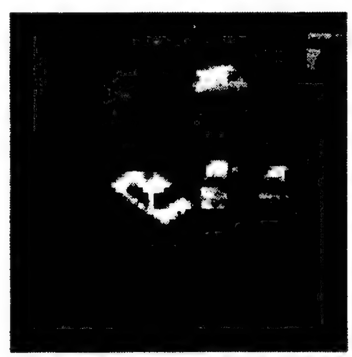
Taxi Sequence Frame 2

Frame 1-2 Optical Flow X Component

Frame 1-2 Optical Flow Y Component



Reconstructed Taxi Frame 1



Reconstructed Taxi Frame 2

Figure 6-4 Taxi Sequence

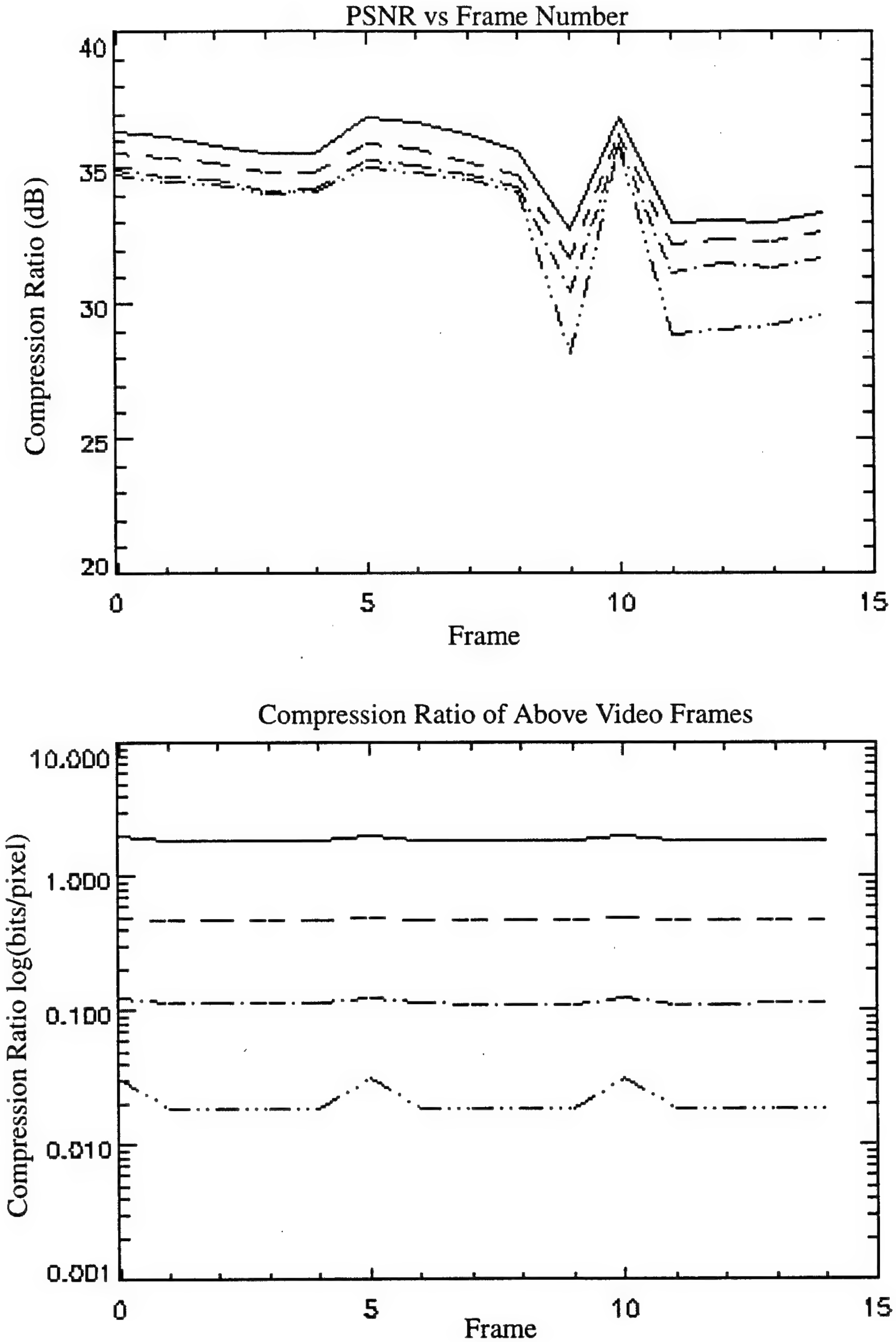


Figure 6-5 PSNR of Video Compression at Various Compression Ratios

Chapter 7

Conclusion

The fractal-wavelet method offers a significant improvement over existing techniques because of its unified approach to image analysis and compression. The fractal wavelet method itself gives naturally higher compression and better reproductive quality than conventional DCT-based methods. By its wavelet frequency division process, it gives a more natural organization to existing fractal methods and allows more accurate block matching. As a result of

its modulus maxima shape representation it gives a shape to texture content-based approach to compressed file organization. By its gradient based block matching technique it is significantly faster than existing wavelet-fractal compression methods. Because of its progressive multiresolution frequency decomposition method it gives faster and more efficient means of image decompression than conventional shape-based wavelet techniques. And by efficient organization of content via shape, texture, color, and motion it produces an efficient object-based video compression method.

The preceding wavelet-fractal method for image compression offers a new framework for data compression and analysis. This model will hopefully serve as a guide both to future research as well as provide a useful technique for image and video compression for immediate needs. It holds the promise for even higher compression ratios with greater spatial resolution and increased speed. Also, because of the well known wavlet fractal signal modelling of natural phenonmena, it offers the possibility of linking physical modeling to existing compression techniques. Thus, the wavelet fractal model should benefit researchers in a wide range of disciplines.

Chapter 8

Appendix

Contents

A. Image and Video Analysis Network System Overview

With the unique capabilities of the wavelet fractal video method the technique lends itself well to an application on the existing World Wide Web. The high compression capability coupled with the flexibility of object based video and pattern matching are uniquely suited to a Web-based system. Figure A-1 shows an over view of how such a system would be organized on the web.

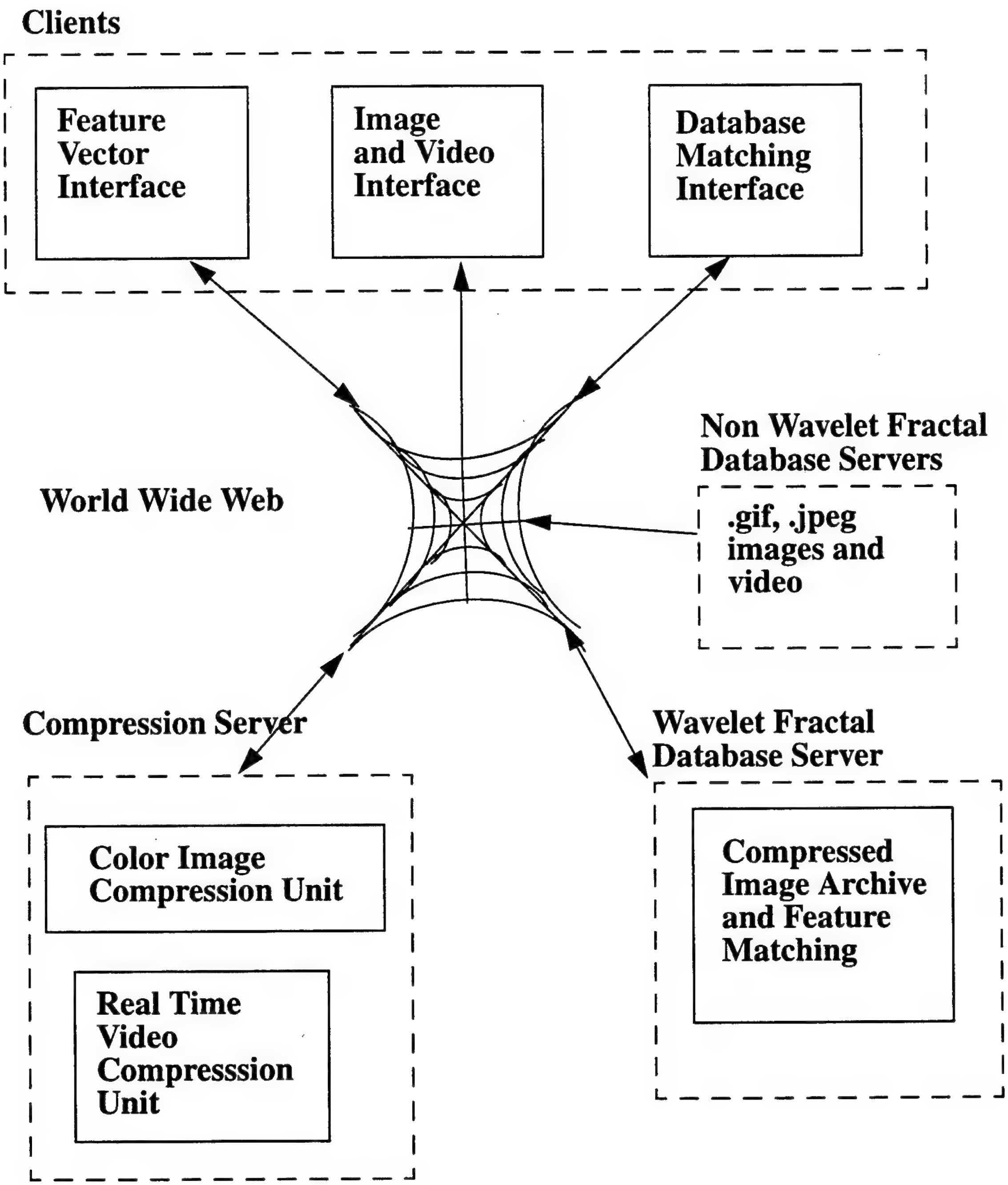
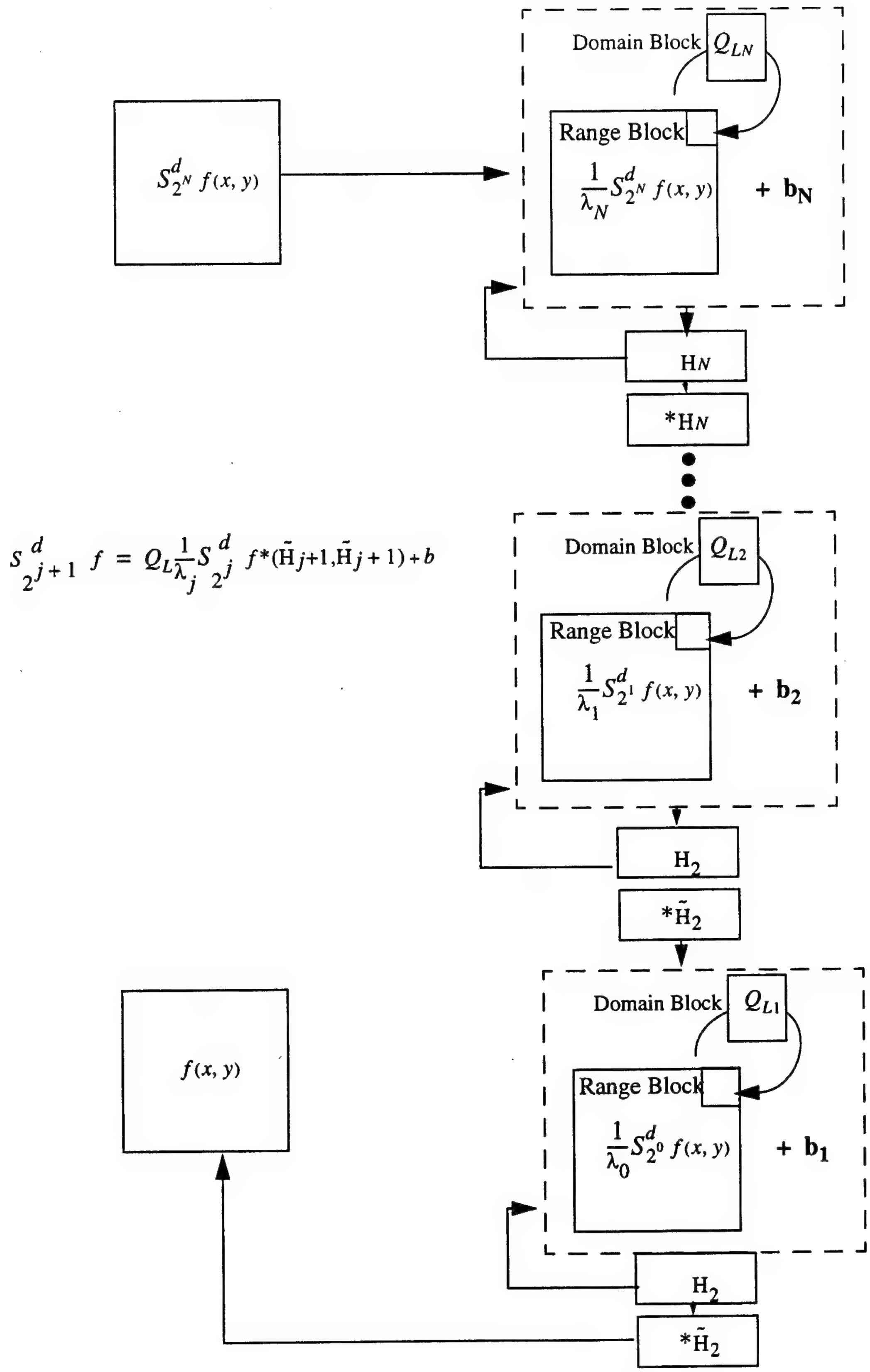


Figure A-1 Network Wavelet Fractal System

A.1 Clients

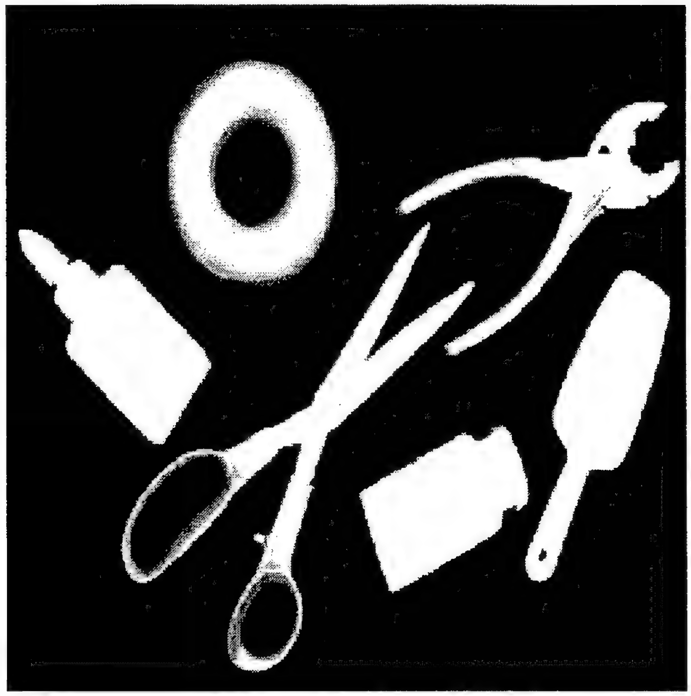
There are essentially three client applications - these applications are exported to the user on run time of a Web browser currently through the Java/HTML programming languages. Once established on the client these applications fulfill three roles. The feature extraction software provides a flexible interface for users to read in an image from any server on the web and convert it into the elemental feature vectors of shape, color, and texture as is done in the preprocessing steps of the video. In this way , various objects in any image format such as GIF and JPEG can be matched against compressed images in the wavelet fractal database for comparison. Secondly the image and video display allows users to display compressed imagery and video in complete or object form either from stored files or live imagery. The decompression technique for images is shown in Figure A-2.



A-2 Direct Reconstruction Technique

Finally the pattern matching interface allows the user to display individual objects in the video scene and match them against shape color and texture vectors of objects in the compressed video database interactively - controlling the relative weighting of the shape, color and texture matching vectors Objects may be removed or added as is shown Figure A- 3.

Original Objects



Self Encoded Objects

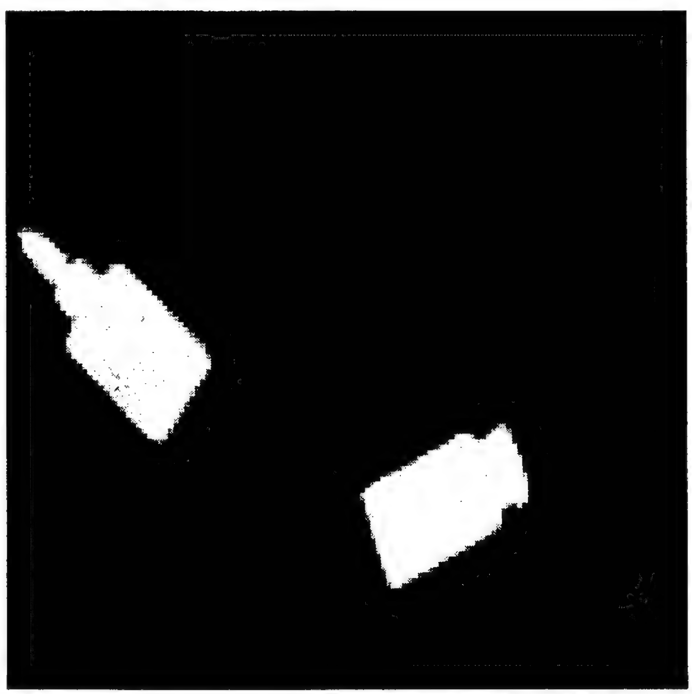
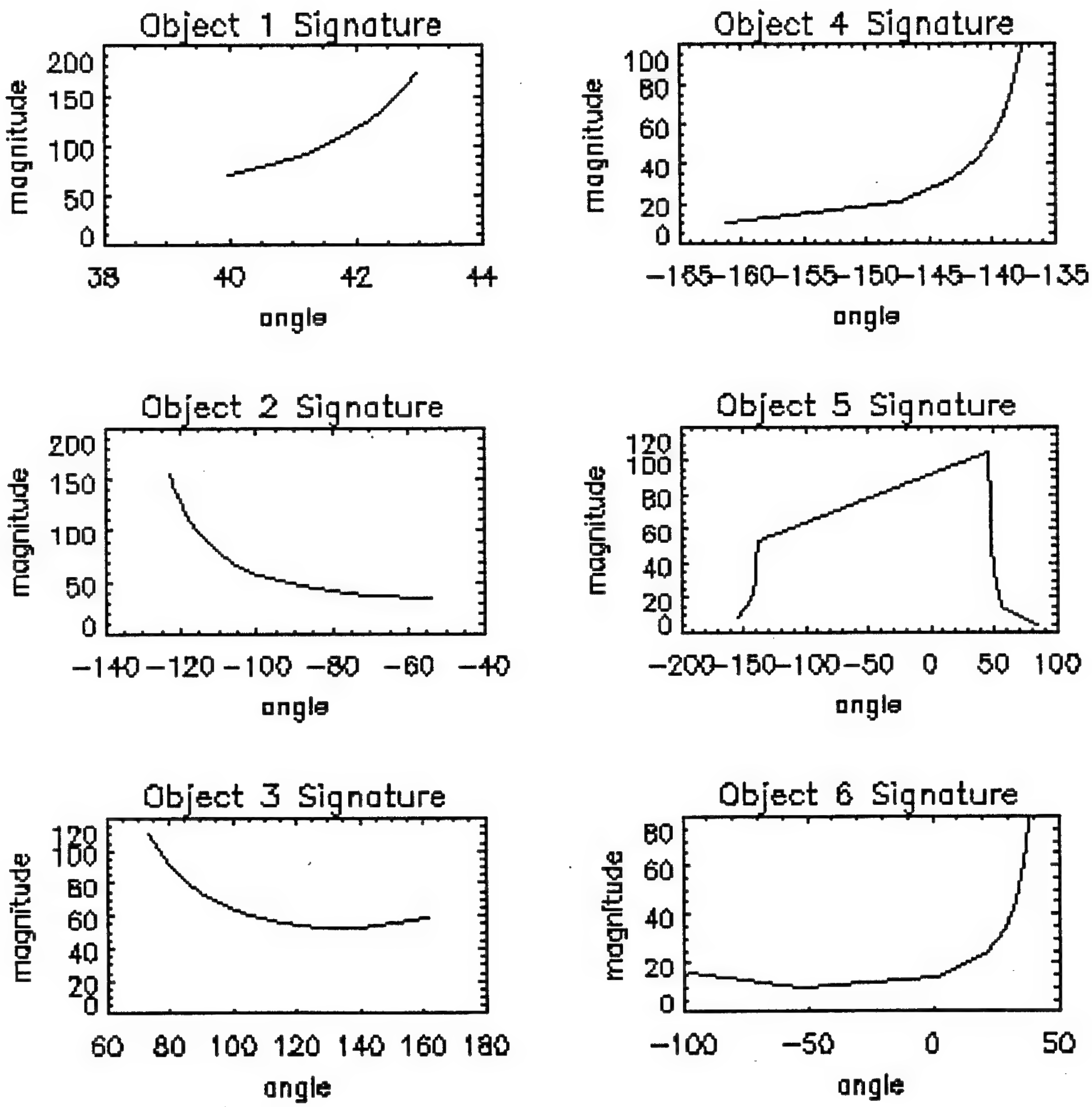
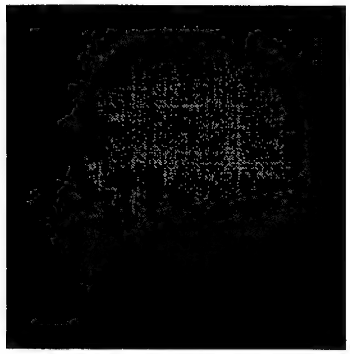


Figure A-3 Self Encoded Object Image

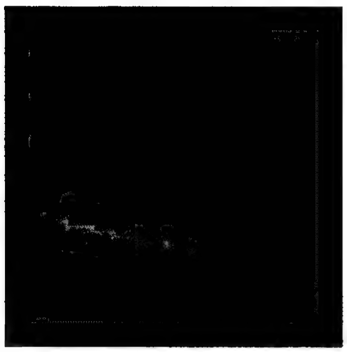
Once the desired objects are selected shape information such as the signatures of Figure A-4 are used to match against the compressed database information at the wavelet fractal database server which will be described later. Additionally the database object matching interface may create composite images either automatically or with user assistance as is shown in Figure A-5



A-4 Signatures of Objects in Figure A-3



Original Image



Object 1

Object 2

Composite Object

A-5 Image Object Composites

A.2 Servers

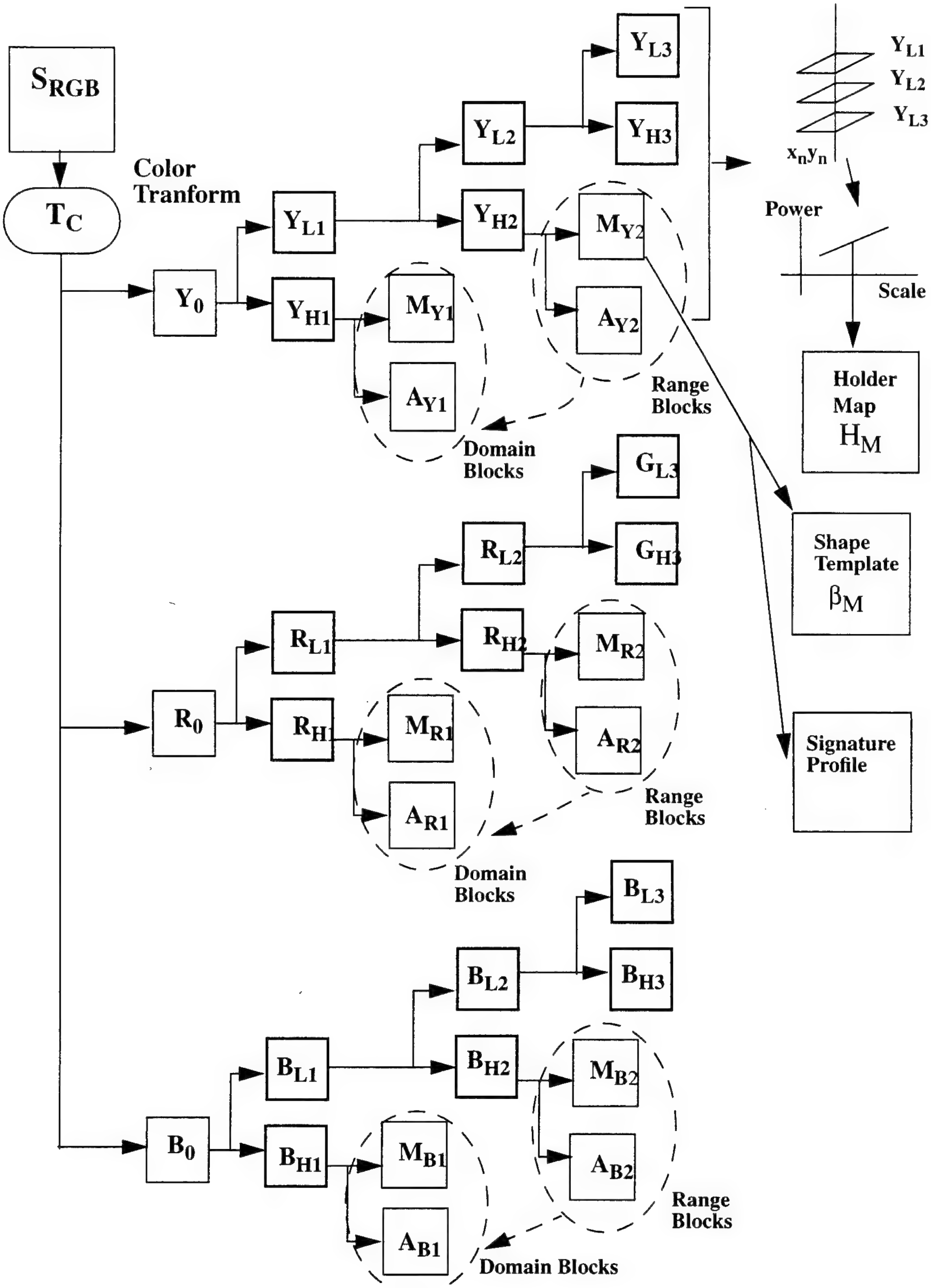
Like the clients there are three servers. Each server is suited to its particular video task. For instance the image and video compression server usually consists of a video camera with special purpose hardware designed to transmit video to the Java Client. The Wavelet Fractal database sever on the other hand is usually a large multiuser computer sometimes with many processing nodes designed for feature vector query.

A2.1 Image and Video Compression Server

The image and video compression server is designed to transmit java software to clients and to compress imagery and video. In many cases this operational can be contained in one portable PC so that many sources of imagery and video may be provide live from multiple locations to the client software.

A2.1.1 Image Compression

The color image compression operation is performed on the server as shown in Figure A-5. The basic color compression compresses Y, R, and B color planes separately as shown in a multiresolution framework. The resulting image file is shown in Figure A-7 This content indexable file is easily referenced by the database matching system.



A-6 Graphical Overview of Color Image Compression

Luminance (Y) Image

Object 1

Header $\overline{\alpha}_{MO1}$ V_{CO1} CentroidX CentroidY 2nd Mom Ang #Range Blocks

Scale 0

Domain X	Domain Y	Range X	Range Y	R_L	b	Optional H
10	10	15	17	2	36	-.2
15	20	8	12	4	5	-.1
25	20	14	33	1	-15	0.4
.	.	.	.	.	.	

Range Blocks - Edge Blocks First

Scale 1
Scale 2
.
.
Scale N

Object 2

Object 3 (background)

$S^d_{2^N} f(x, y)$ ← **lowpass image**

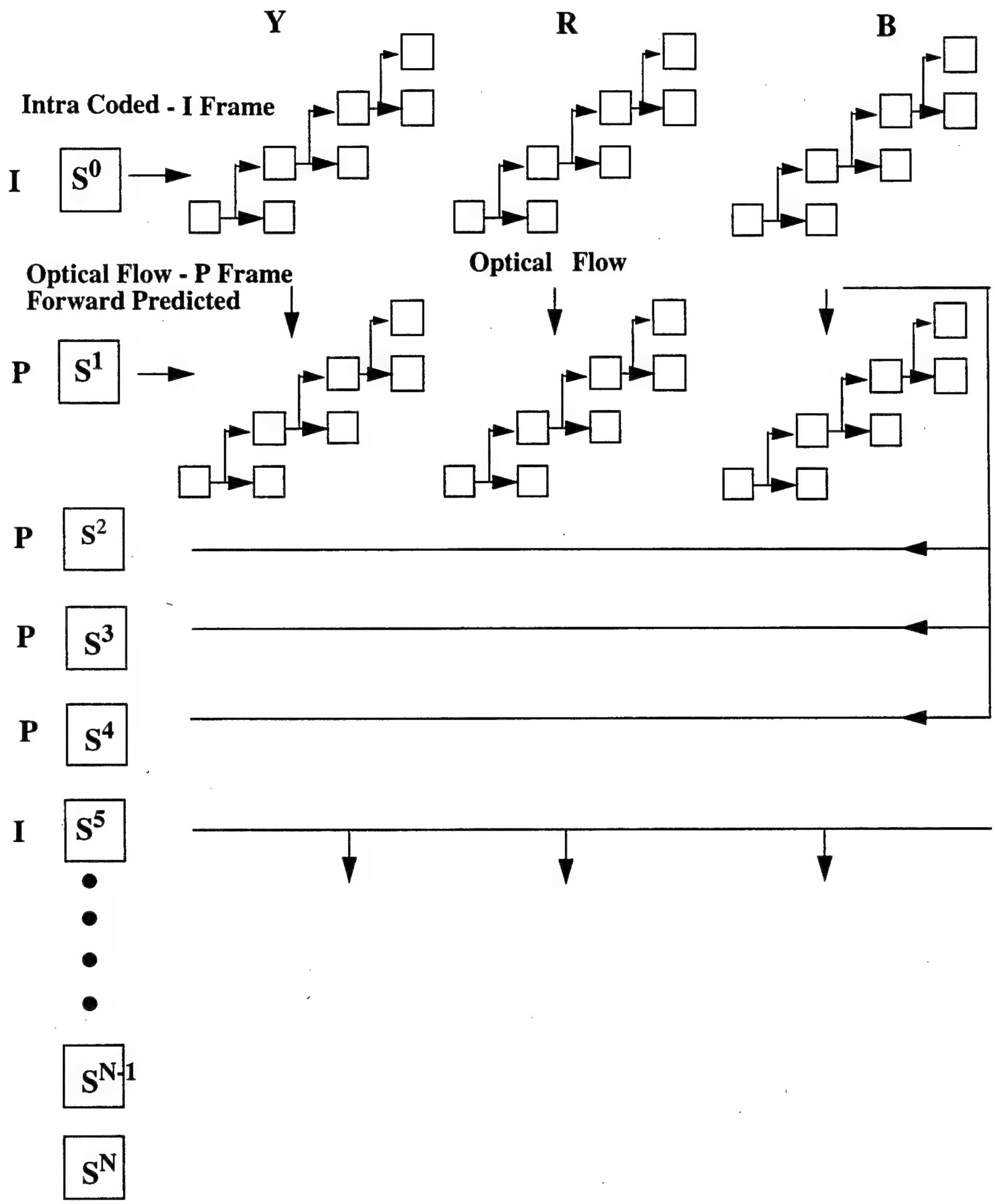
Red (R) Image - I Frame

Blue (B) Image - I Frame

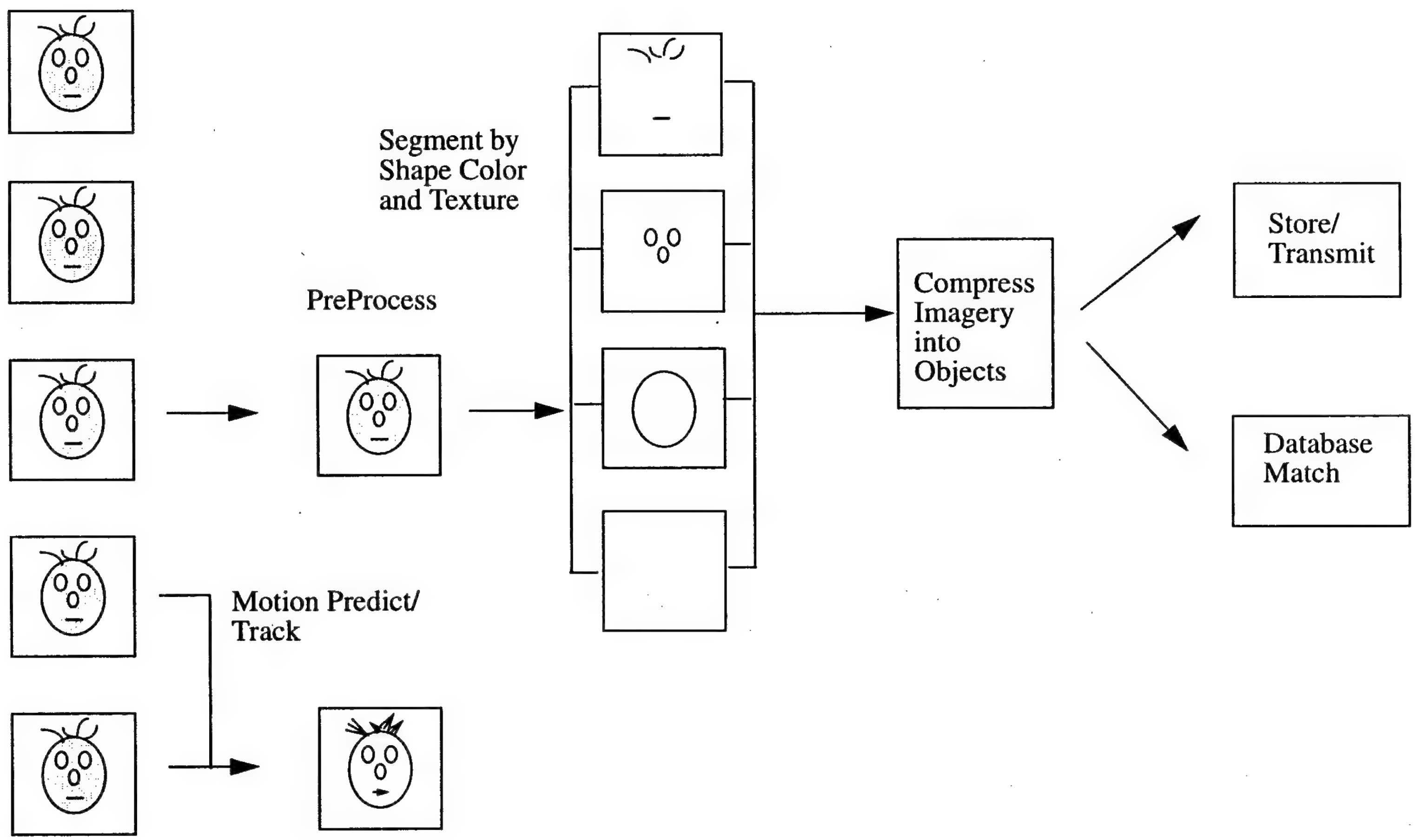
A-7 Compressed Image File Format

A2.1.2 Video Compression System

The image and video server also contains the video compression system. This the video compression process is depicted in Figure A-8. The Y,R, and B fields are computed separately and the optical flow between these fields is used to determine what compressed blocks to transmit. This function be performed non real time or real time with a camera interface and special purpose hardware. A-9 shows how a typical video sequence is broken into its component shape color and texture fields for content based query and object video. This process is then illustrated with actual imagery in figures A-10 through A-11. Finally in Figure A.-12 the compressed file format is illustrated. What is clear in the image compression process is that there is a great deal of flexibility with the wavelet fractal technique in terms of adapting to a specific network bandwidth. The spatial, temporal, object, and color scalable features of this compression system make it practical for use on the low bandwidth World Wide Web.



A-8 Graphical Overview of Video Compression



A-9 Object Oriented Video Coding

Luminance Image Frame 1

Luminance Image Frame 2

Object Image

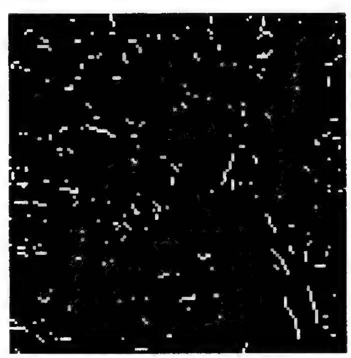
Texture Image

Red Image

Blue Image

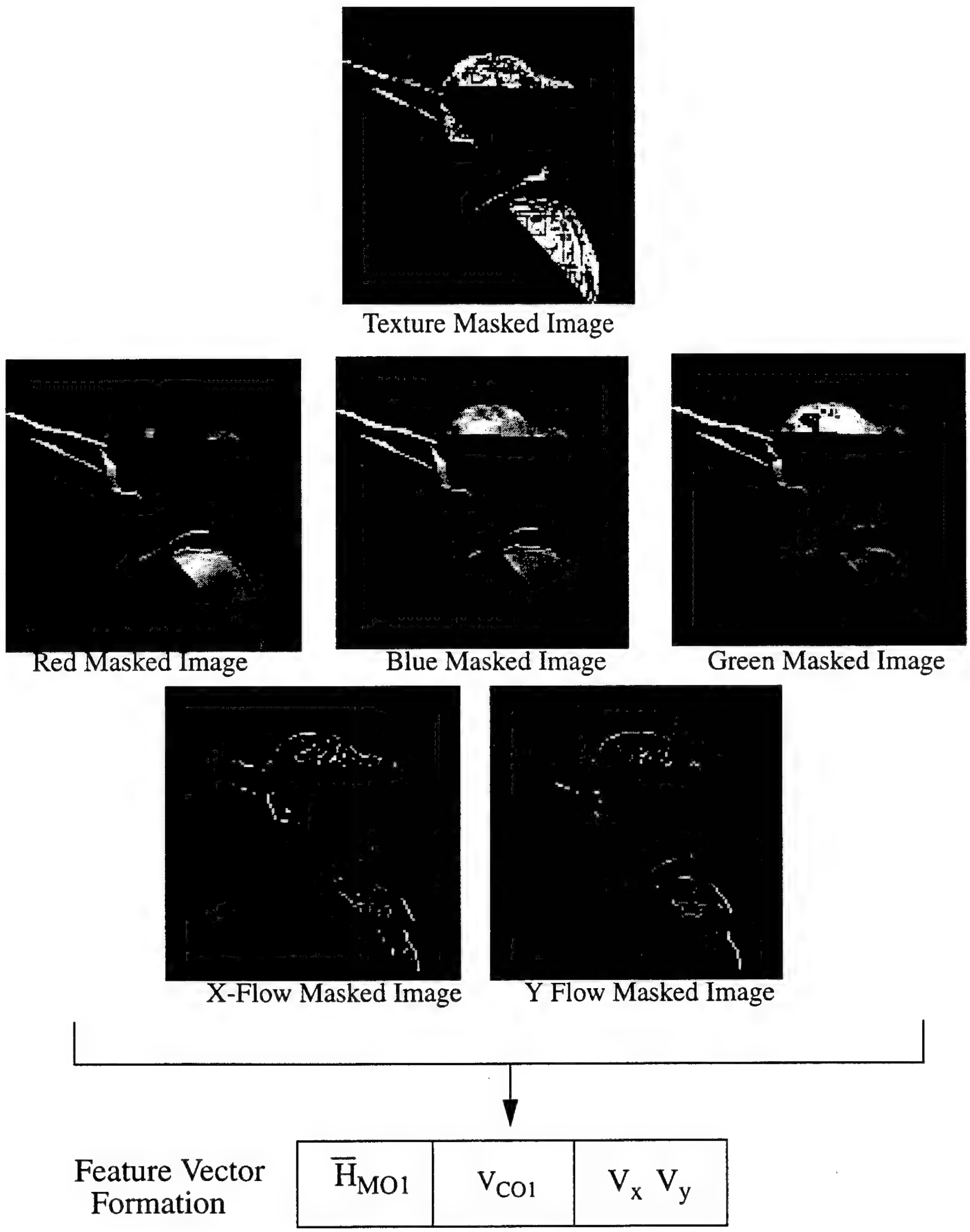
Green Image

X Optical Flow



Y Optical Flow

A-10 Video Image Segmentation



D-3 Video Object Creation

Luminance (Y) Image - I Frame

Object 1

Header $\overline{\alpha}_{MO1}$ V_{CO1} CentroidX CentroidY 2nd Mom Ang #Range Blocks

Scale 0

Domain X	Domain Y	Range X	Range Y	R_L	b	Optional H	V_x	V_y
10	10	15	17	2	36	-.2	5.3	2.1
15	20	8	12	4	5	-.1	2.2	2.3
25	20	14	33	1	-15	0.4	1.1	1.5
.	.	.	.	.	.			

Range Blocks - Edge Blocks First

Scale 1
Scale 2
.
.
Scale N

Object 2

Object 3 (background)

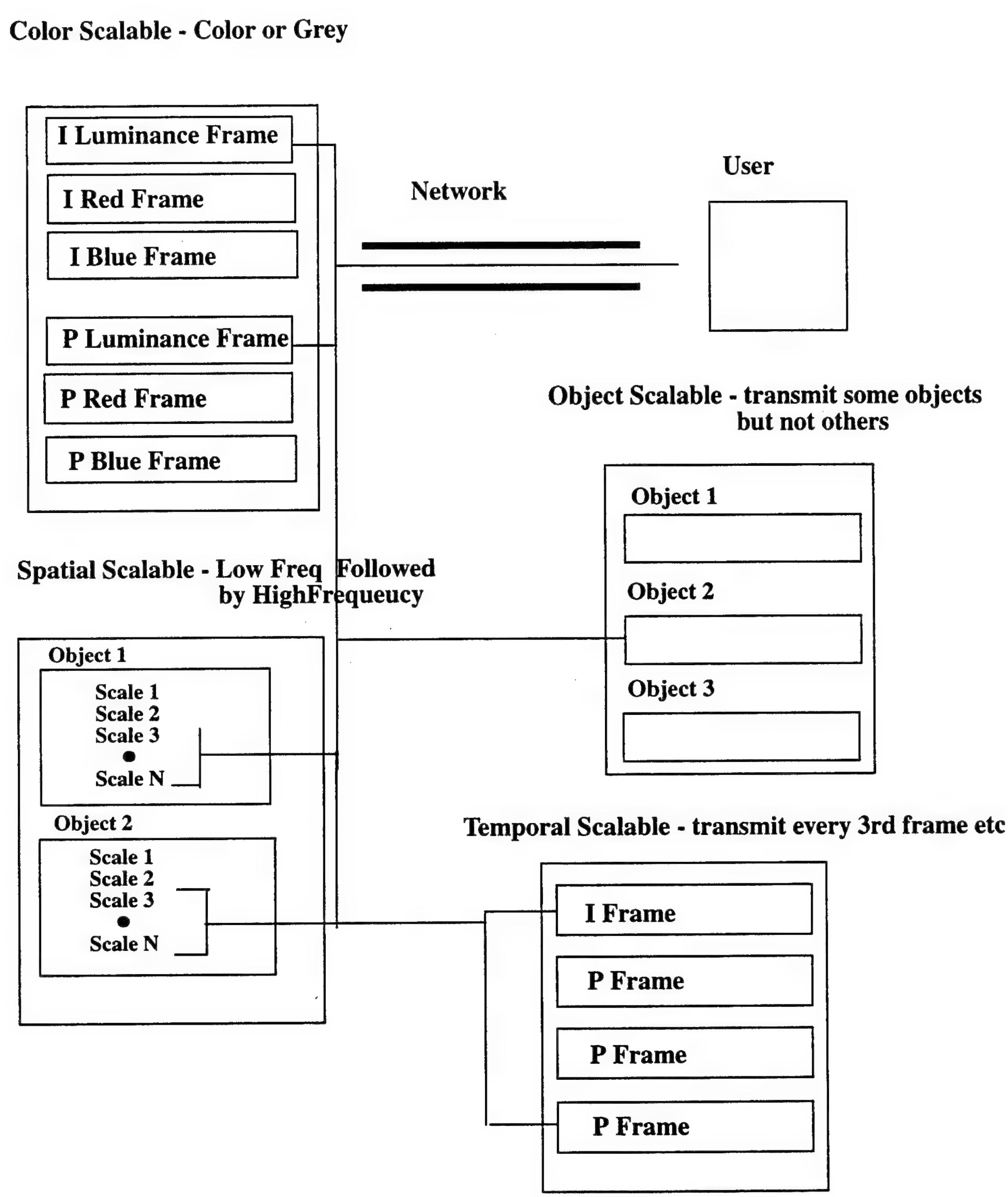
$S^d_{2^N} f(x, y)$ ← **lowpass image**

Red (R) Image - I Frame

Blue (B) Image - I Frame

P Frames (if video)

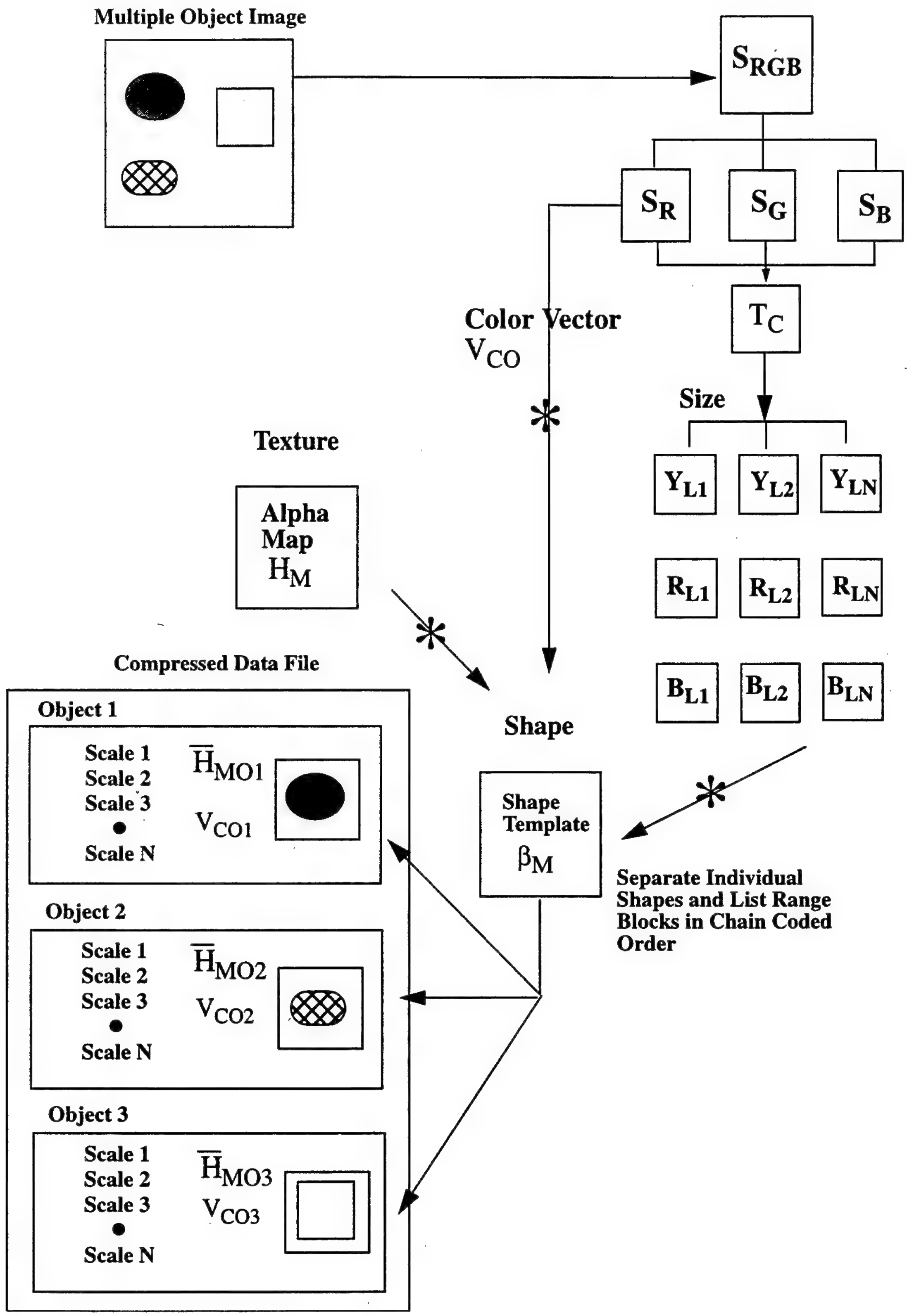
A-12 Compressed Object Video File



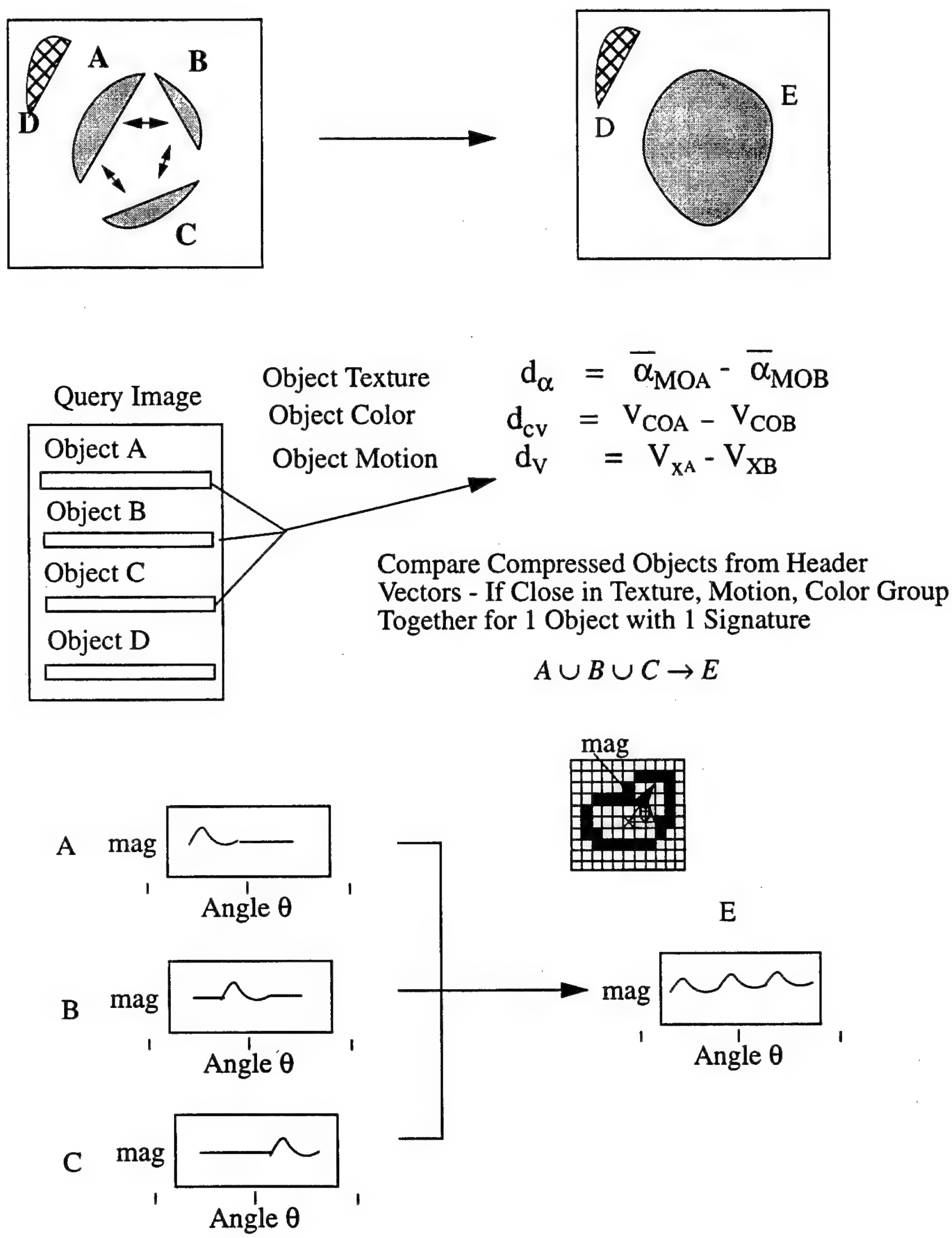
A-13 Video Scalability

A3 Wavelet Fractal Database Server

The wavelet fractal database server servers two roles. First it is a repository of data in the wavlet fractal compressed data format. Secondly, it performs searches across vast numbers of compressed files very efficiently for object and image matches. This role is nicely filled by some of the large supercomputer architectures with multiple processors, large network capacity, and large disk capacity. The process of creating and submitting a query with either the feature or database matching interfaces is shown in Figures A-14. and A15 . Many methods for matching vectors may be considered including neural networks.



A-14 Compressed Domain Object Matching



A-15 Object and Signature Compositing

B Object Moments

Simple shapes can be characterized by zeroth, first, and second moments. Zeroth moment is the area in pixels occupied by an object. If I(x,y) is a luminance image with pixels of coordinates x,y then the zeroth moment is.

$$M0 = \sum_{x=0}^{N} \sum_{y=0}^{N} I(x, y) \tag{8.1}$$

First moment is the intensity weighted average in the x and y directions of an object also known as the center of mass or centroid of an object. This information is necessary for determining signature profile and can also be used to track objects in motion video as is discussed in the next chapter.

$$M1x = \frac{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} I(x, y) \cdot x}{M0} \tag{8.2}$$

$$M1y = \frac{\sum_{x=0}^{N-1} \sum_{y=0}^{N-1} I(x, y) \cdot y}{M0} \tag{8.3}$$

Finally, second moment gives the angle θ of the primary axis of orientation of the object. This information can be very useful in the signature matching process for lining up signatures of the same objects that has been rotated.

$$M2a = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} (M1x - x)^2 I(x, y) \tag{8.4}$$

$$M2b = 2\sum_{x=0}^{N-1}\sum_{y=0}^{N-1}(M1x-x)(M1y-y)I(x,y) \tag{8.5}$$

$$M2c = \sum_{x=0}^{N-1}\sum_{y=0}^{N-1}(M1y-y)^2 I(x,y) \tag{8.6}$$

$$\tan 2\theta = \frac{M2b}{M2a-M2c} \tag{8.7}$$

Bibliography

Contents

1 A. Arneodo, F. Argoul, J.F. Muzy and M.Tabard, “Beyond Classical Multifractal Analysis Using Wavelets: Uncovering a Multiplicative Process Hidden in the Gometrical Complexy of Diffusion Limited Aggregates”, Fractals, Vol 1, No. 3, pp 629-649, 1993.

2 Alexandru Bogdan, "Multiscale (Inter/Intra-Frame) Fractal Video Coding", 1st IEEE International Conference on Image Processing (ICIP-94), Austin, TX, November 1994

3 Alexandru Bogdan, "The Fractal Pyramid with Applications to Image Coding", IEEE International Conference on Acoustics, Speech & Signal Processing (ICASSP-95), Detroit MI, May 1995.

4 Alexandru Bogdan, Image Coding Using Iterative Transformations with Applications to Image Communication, PhD Thesis, Columbia University, Feburary 1995.

5 A. Aydin Alatan, L. Onural, IEEE International Conference on Image Processing Proceedings of the 1995 IEEE International Conference on Image Processing, Part 1 of 3 Oct 23-26 1995 v1 1996 Washington DC.

6 R. Bonneau, "Image Processing Techniques in Automatic Target Tracking", AIAA Fire Control Symposium, Boulder Colorado, July 1992.

7 R.Bonneau, "Using the centroid operator for faster multiresolution image compression and pattern recognition", SPIE Conference on Mathematical Imaging: Wavelet Applications in Signal and Image Processing, San Diego, July 1995.

8 R. Bonneau, "The multiresolution transform and its application to image coding", SPIE/IEEE Conference on Visual Communication and Image Processing: Wavelets and Fractals, Orlando, March 1996.

9 R. Bonneau, "The multiresolution transform and its application and its application to shape recognition in the compressed domain", SPIE AeroSense 96: Hybrid Image and Signal Processing V, Orlando, April 1996.

10 R. Bonneau, "A multiresolution transform and its application to video coding ", SPIE/ EUROPTO, European Symposium on Imaging and NetworkTechnologies, Berlin,October 1996.

11 M. Barnesley, Fractals Everywhere 2nd Ed., Academic Press, New York, 1993.

12 M. Barnsely and L. Hurd, *Fractal Image Compression*, A.K. Peters, Wellesley Mass., 1993.

13 R. Carmona, W. Hwang, B. Torresani, "Characterization of Signals by the Ridges of Their Wavelet Transforms", UC Irvine Technical Report, March 1995.

14 J. Canny, "A Computational Approach to Edge Detection", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol PAMI-8, No. 6, November 1996.

15 P. Charton, V. Perrier, "A Pseudo-Wavelet Scheme for the Two-Dimensional Navier-Stokes Equations", Matematica Aplicada e Computacional, March 1996

16 P. Cheng, J. Li, and C. Kuo, "Mulitscale Video Compression"Using the Wavelet Transform and Motion Compensation", IEEE International Conference on Image Processing, Washington DC, Oct 1995.

17 C.T. Chu, D. Anastassiou, SF Chang, To appear in Signal Processing:Special Issue on MPEG4, "Hybrid Object-Based/Block-Based Coding in Video Compression at Very Low Bit Rate"

18 A. Cohen, I. Daubchies, J.C. Feauveau, "A Biorthogonal Bases of Compactly Supported Wavelets", Cummunications on Pure and Applied Mathematics, Vol. XLV, 485-560, 1992.

19 A. Davis, A. Marshak, and W. Wiscombe, "Wavelet Based Mulitfractal Analysis of Non-Stationary and/or Intermittent Geophysical Signals", Wavelets in Geophyics, Academic Press, pp.

249-298.

20 G. Davis, "Adaptive Self-Quantization of Wavelet Subtrees: A Wavelet-Based Theory of Fractal Image Compression", SPIE Conference on Mathematical Imaging: Wavelet Applications in Signal and Image Processing, San Diego, July 1995.

21 I. Daubechies, "Ten lectures on wavelets," CBMS-NSF Series Appl. Math. SIAM, 1991.

22 W. Dahman, A. Kunoth, K. Urban, " A Wavelet-Galerkin Method for the Stokes Equations", Institute fur Geometrie und Prakicshe Mathematik internal report, August 1995.

23 J. Elton, "An Ergodic Theorem for Iterated Maps", Ergodic Theory and Dynamical Systems, Number 7, pp. 481-488, 1987.

24 C. Eversz, K. Berkner, and W. Berghorn, "A Local Multiscale Characterization of Edges Applying the Wavelet Transform", Proceedings of NATO ASI, Fractal Image Coding and Analysis, Trodheim Norway, July 1995.

25 M. Farge., "Wavelet Transforms and Their Application to Turbulence", Annual Review of Fluid Mechanics, Vol 24, pp. 395-457, 1992.

26 Y. Ficher, *Fractal Image Compression*, Springer Verlag, New York, 1995.

27 Y. Fischer, Rogvin, T.P. Shen, "Fractal (Self VQ) Encoding of Video Sequences", Proceedings of the SPIE, Visual Communications and Image Processing, 25-28 September 1994.

28 P. Flandrin, "Wavelet analysis and synthesis of fractional Brownian motion", IEEE Transactions on Information Theory, vol. 38, no. 2, March 1992.

29 U. Frisch, P. Sulem, and M. Nelkin, "A Simple Dyamical Model of Intermittent Fully Developed Turbulence", Journal of Fluid Mechanics, Vol. 87, Part 4, pp. 719-736, 1978.

30 D. Giusto and Stefano Fioravanti, "Estimation of qth-Order Fractal Dimensions", ICASSP, IEEE International Conference on Acoustics, Speech, and Signal Processing, Adelaide Austraila,

Part 5, pp. 361-364.

31 D. Gorsich and Charles Tolle, "Wavelet and Fractal Analysis of Ground Vehicle Images", SPIE Wavelet Applications in Signal and Image Processing IV, Dever Colorado, pp 109-119, August 1996.

32 R. Gonzales and P. Wintz, *Digital Image Processing 2nd Ed.*, Addison Wesley, Reading Mass., 1987.

33 C. Herley and M. Vetterli. "Wavelets and Filter Banks: Theory and Design." IEEE Transactions on Signal Processing, 40(9), September 1992.

34 M. Hotter, "Optimization and Efficiency of an Object-Oriented Analysis-Synthesis Coder", IEEE Transactions on Circuits and Systems for Video Technology, Vol 4., No 3, June 1994.

35 D. Hubel, *Eye Brain and Vision*, Scientific American, New York, 1988.

36 W.H. Hwang and Stephane Mallat, "Characterization of Self-Similar Mulitfractals with Wavlet Maxima", NYU Technical Report 641, July 1993.

37 A. Jacquin, A Fractal Theory of Iterated Markov Operators With Applications to Digital Image Coding, Georgia Institute of Technology PhD. Dissertation, August 1989.

38 A.K. Jain, Fundamentals of Digital Image Processing, pp. 382-383, Prentice Hall, Englewood NJ, 1989.

39 L.M. Kaplan and C.C. Kuo, "Texture Segmentation via Haar Fractal Feature Estimation", Journal of Visual Communication and Image Representation, Vol 6, No 4, December, pp 387-400, 1995.

40 H. Krupnik, D. Malah and E. Karnin, "Fractal Representation of Images via the Discrete Wavelet Transform", IEEE 18th Conv. of EE in Israel, Tel-Aviv, March1995.

41 H. Kim and C. Li, "A Non-orthogonal Wavelet Edge Detector with Four Filter Coeficients",

SPIE Mathematical Imaging Vol 2034, pp. 115-126, 1993.'

42 S. Lowen and M. Teich, "Estimation and Simulation of Fractal Stochastic Point Processes", Fractals Vol 3, No 1 pp. 183-210, 1995.

43 Y. Lu, "Reasoning about Edges in Scale Space", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol 14, No 4, April 1992.

44 S. Mallat and S. Zhong, "Characterization of signals from multiscale edges", IEEE Trans. on Patt. Anal. and Mach. Intell., Vol. 14, pp.710-732, July 1992.

45 S. Mallat, "Multiresolution Approximations and Wavelet Orthonormal Bases of L2R", Transactions of the American Mathematicall Society, Volume 315, Number 1, September 1989.

46 B. Mandelbrot, *The Fractal Geometry of Nature,* Freeman, San Francisco, CA 1983.

47 R. Mehotra and J. Gary, "Similar-Shape Retrieval In Shape Data Management", IEEE Computer, pp. 57-62, September 1995.

48 B. Moghaddam and A. Pentland, "Probablilistic Visual Learning for Object Detection", 5th International Conference on Computer Vision, Cambridge, MA, June 1995.

49 H. Murase and S. Nayar, "Visual Learning of 3-D Objects From Appearance", International Journal of Computer Vision, January, 1995.

50 J.F Muzy, E. Bacry, A. Arneodo, "Multifractal formalism for fractal signals: The structure function approach versus the wavelet transform modulus maxima method.", Physical Review E. Volume 47, Number 2, Februrary 1993.

51 J.F. Muzy, E. Bacry, and A. Arneodo, "The Multifractal Formalism Revisited with Wavelets", International Journal of Bifurcation and Chaos, Vol 4., No 2, (1994) pp. 245-302.

52 A. Netravali and B. Haskell, *Digital Pictures Representation, Compression, and Standards*, Plenum, New York, 1995.

53 H. Peitgen, H. Jurgens, D. Saupe, *Chaos and Fractals*, Springer Verlag, New York, 1992.

54 A. Pentland, “Fractal Based Description of Natural Scenes”, IEEE Transactions on Pattern Analysis, and Machine Intelligence, Vol PAMI-6, No 6., November 1984.

55 R. Reidi and B. Mandelbrot, “Inversion Formulas for Continuous Multifractals”, Advances in Mathematics Preprint, July 1995.

56 R. Reidi and J. Vehel, “Multifractal Properties of TCP Traffic”, IEEE/ACM Transacations on Networking, January 1997.

57 R. Rinaldo and G. Calvagno, "Image coding by block prediction of multiresolution subimages", IEEE Transactions on Image Processing, Vol. 4, No. 5, May 1995, pp. 682-687.

58 C. Schwartz, G. Baum, E.N. Ribak, “Turbulence Degraded Wavefronts as Fractal Surfaces”, J. Optical Society of America, Vol 11, No 1, January 1994.

59 J. Shapiro, “Embedded Image Coding Using Zerotrees of Wavelet Coefficients”, IEEE Transactions on Signal Processing, Vol 41, No 12, December 1991.

60 Y. Shoham and A. Gersho, “Efficient Bit Allocation for an Arbitrary Set of Quantizers”, IEEE Transactions on Acoustics, Speech and Signal Processing, Vol 36, No. 9 September 1988.

61 Sim D.G., Park, R.H, “Two stage algorithm for motion discontinuity-preserving optical-flow-estimation”, Computer Vision and Image Understanding v65 n1 Jan 1997, Academic Press Inc, San Diego CA USA p 19-37.

62 J. Smith and S. Chang, "Frequency and Spatially Adaptive Wavelet Packets", IEEE International Conference on Acoustics, Speech & Signal Processing (ICASSP-95), Detroit, May 1995.

63 J. Smith and S.F. Chang, " Local Color and Texture Extraction and Spatial Query", IEEE International Conference on Image Processing, 1996.

64 J. Smith and S.F. Chang, "Quad Tree Segmentation for Texture-Based Image Query", ACM

2nd International Conference on Mulitmedia, SanFrancisco CA, Oct 1994.

65 Z. Struzik, "The Wavelet Transform in the Solution to the Inverse Fractal Problem", Fractals, Vol. 3, No. 2, 329-350, 1995.

66 M. Sun, C. Li, R. Sclabassi, "Edge Localization in Images by Symmetrical Wavelet Transforms", SPIE Mathematical Imaging, Vol 2034, pp. 92-99, 1993.

67 V. Sundeswaran., Global Methods for Image Motion Analysis, Ph. D Thesis, Computer Science Department, New York University, September 1992.

68 R.I. Sykes and R.S. Gabruk, "Fractal Representation of Turbulent Dispersing Plumes", Journal of Applied Metorology, Vol 33. pp. 721-732, June 1994.

69 B. Tseng and D. Anastassiou, "A Theoretical Study on an Accurate Reconstruction of Multiview Images based on the Viterbi Algorithm. In International Conference on Image Processing", ICIP '95, Washington, D.C., October 1995.

70 D. Valentin et.al., "Connectionist Models of Face Processing: A Survey", Pattern Recognition, Vol 27, No 9., pp1209-1230, 1994.

71 M. Vetterli and C. Herley, "Wavelets and Filter Banks: Theory and Design", IEEE Transactions on Signal Processing, Vol. 40 No 9, September 1992.

72 M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*, Prentice Hall, Englewood New Jersey, 1995.

73 B. Vidakovic and P Muller, "Wavelets for Kids", Duke University internal report number 42A06.

74 J. Villasenor, B. Belzer, J. Liao, "Wavelet Filter Evaluation for Image Compression", IEEE Transactions on Image Processing Vol. 4, No 8, August 1995.